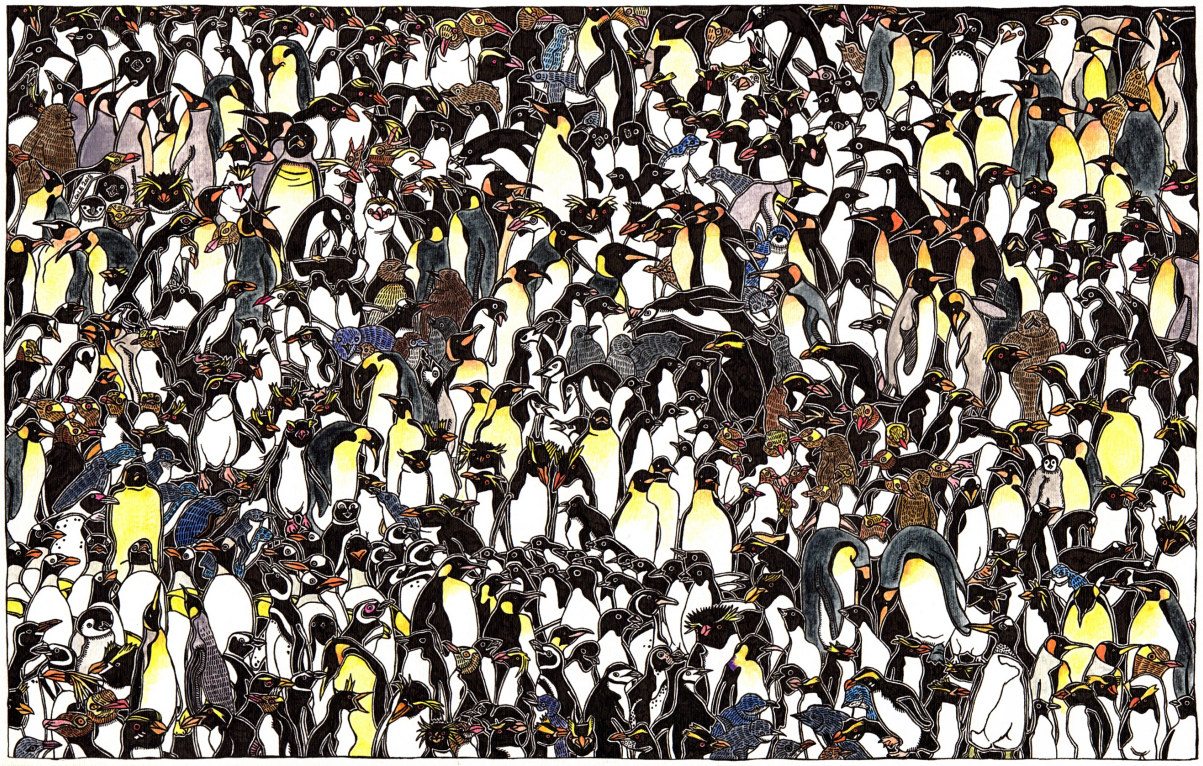
Fire and ice: Genomic insights into the evolution, biogeography and extinction of modern penguins

Theresa L Cole

A thesis submitted for the degree of

Doctor of Philosophy



Department of Zoology

University of Otago

September 2019

Cover art: *Penguins*. Theresa L Cole, 2018.

Abstract

Islands of the vast Southern Ocean host abundant endemic wildlife populations representing important breeding grounds for many seabirds. With contrasting geological, glacial and human-impact histories, these islands represent strong systems for inferring evolutionary processes. Although penguins (Sphenisciformes) spend much of their lives at sea, most taxa require ice-free terrain for breeding, inhabiting every major landmass and archipelago in the Southern Ocean. While penguins are distributed widely across sub-Antarctic and Antarctic coastlines, nearly a third of all taxa are endemic to geologically young islands, especially in the New Zealand (NZ) region. Penguins therefore represent an ideal group of seabirds with which to study the biogeographic and evolutionary effects of island history. This thesis extends classic approaches in island biology research to the Southern Ocean islands, using penguins as the focal system. The various analyses include evidence from near-complete mitochondrial genomes (mitogenomes), thousands of genome-wide single nucleotide polymorphisms (SNPs), together with mitochondrial cytochrome oxidase subunit 1 (COI) and control region (CR) sequences, generated to examine spatial and temporal evolutionary patterns across all extant and recently extinct taxa. First, this thesis generates a broad framework for understanding penguin evolution by providing the first time-calibrated phylogeny to encompass all modern taxa. Divergence-time estimates demonstrate that many island-endemic taxa diverged following the geological emergence of their natal islands. This study also provides the first molecular demographic analyses to encompass all sub-Antarctic and Antarctic penguin taxa. Genome-wide SNPs yielded signatures of concerted demographic expansions following the Last Glacial Maximum (LGM), suggesting taxa inhabiting islands south of the LGM sea-ice limit underwent rapid post-glacial expansions associated with rapid climate change. This result is further supported by very low population structure across entire Southern Ocean distributions. This thesis provides the first multi-species genetic assessment of the impacts of Polynesian and European arrival on NZ and Chatham Island penguin assemblages. Phylogenetic analyses of modern and ancient mitogenomes, together with COI and CR sequences, revealed the presence of two recently extinct penguin taxa from the Chatham Islands: *Eudyptes warhami* and *Megadyptes antipodes richdalei*, that are formally described here using genetic and morphological comparisons. Recent extinction of these island-endemic lineages was likely a direct result of human pressure. By contrast, phylogenetic and demographic analyses of COI and CR sequences revealed limited evidence for demographic reductions in NZ *Eudyptes* penguins, with ecological preferences possibly buffering those taxa

from human impacts. This thesis also provides new systematic insights for 'extinct' prehistoric penguins elsewhere, demonstrating that the so-called Hunter Island penguin is actually an artificial assemblage of three extant taxa. Finally, this study concludes that rockhopper penguins comprise three species (*Eudyptes moseleyi*, *E. filholi* and *E. chrysocome*), that extant and extinct *Megadyptes* penguins comprise three subspecies (*M. antipodes antipodes*, *M. a. waitaha* and *M. a. richdalei*) and that macaroni (*Eudyptes chrysolophus chrysolophus*) and royal (*E. c. schlegeli*) penguins are probably incipient species. Overall, this study provides a global assessment of recent penguin biogeography, evolution and extinction, providing a comprehensive basis for ongoing management and conservation of penguin biodiversity.

Preface

This thesis is a synthesis of work led and undertaken by myself between December 2015 and June 2019. My thesis includes three articles published in international peer-reviewed journals and one that is currently under revision.

My research chapters are the result of many national and international collaborations. My co-authors contributed in many ways, and helped me connect island biology, ornithology, geology, climatology and archaeology with genomics, palaeontology and bioinformatics. In recognition that each research chapter was only made possible by the joint efforts of a team of people, the pronoun 'we' is used throughout those chapters. In addition, I have attempted to give each co-author a representative percentage weighting of their contribution for each of my research chapters.

The costs associated with my project were substantial, and were obtained over several years by different people. The samples studied in my project were collected from some of the most remote places on the planet, and in most cases, are irreplaceable. Their collection involved the efforts by many people. While it is difficult to calculate the exact costs involved in my study, my co-authors and I have instead acknowledged all of the funding and permitting sources, and individually recognised the people that assisted us. I apologise to anyone who may have unintentionally been left out.

My interest in science as a future pathway was sparked much later than most of my peers. Although both my parents are scientists, I actually never considered it as a suitable option for myself. In fact, I distinctly remember telling my high school teacher - Mr Jones, that I had no intentions of ever taking a biology class again. I was determined to be a visual artist. Inspired by the biodiversity that surrounded me in the Adelaide Hills, I started exhibiting and selling my art, and started a Bachelor of Visual Arts. To extend my skills (e.g. the cover art), I briefly studied printmaking, glassblowing and silversmithing, and was exposed to chemistry. This is when I realised I needed to know more about our world and how it works. Taking a big risk, I left my familiar surroundings and dove into the unknown, beginning a Bachelor of Science degree. While science and art are often considered opposites, I have found more similarities than differences. Science depends on creativity; in developing new questions, troubleshooting problematic laboratory protocols, creating figures, organising manuscripts and ensuring a balanced life.

Acknowledgements

Without the generous collaboration of so many people, my project would never have evolved into the penguin-wide and circumpolar thesis that it has become. The following five pages is my attempt to thank everyone for the time, support and encouragement that was given to me, and recognise the new friends and collaborations I have made along the way.

I am incredibly lucky to have been supervised by Jon Waters and Janet Wilmshurst. I am especially grateful that you both agreed to support me, as I diverted from what would have been a very different life path. You made me feel welcome and valued in the New Zealand science community.

I thank Jon for being an exceptional primary supervisor and mentor. You were encouraging, supportive, exceptionally fast at providing feedback, and gave me considered and helpful guidance in unfamiliar situations. I am grateful to the freedom you gave me to develop my research, and your confidence that I would get the job done. It is a pleasure to work with you.

I am grateful to Tom Hart, who had intended to extend his recent penguin population genomic research to crested penguins, and had already collected thousands of crested penguin samples before I started. Tom was extremely generous, sent me all the samples, sunk a considerable amount of his own funding into the sequencing, and gave me the freedom to take the project in my own direction. Without this, my study would have been primarily restricted to New Zealand samples. I also thank Tom, Gemma Clucas and Jane Younger, who generated the population genomic datasets for Adélie, gentoo, chinstrap, king and emperor penguins. Without those datasets, my population study would have been restricted to just crested penguins.

I thank Nic Rawlence, Paul Scofield, Alan Tennyson, Jamie Wood, Kieren Mitchell and Alan Cooper for kick-starting the prehistoric parts of my project. Nic had already sampled and extracted DNA from hundreds of New Zealand crested penguin specimens for his own projects, and Jamie and Kieren had already started genomic work on the Chatham Islands penguin material. I am grateful that you all let me lead those projects as part of my thesis, and I appreciate that you gave me the freedom to combine those projects and take them in my own direction. I am also grateful to Kieren and Pere Bover, who taught me how to prepare enriched ancient mitochondrial DNA libraries at the Australian Centre of Ancient DNA, and who also prepared a second DNA library and provided much-needed bioinformatics guidance.

I am grateful to my mentor and friend, Alan Tennyson. Alan went above and beyond the normal role of a collaborator, devoting a lot of his own time to my projects. I was particularly impressed

by Alan's careful attention to detail, and his meticulous measurements of hundreds of penguin bones for this project. I also thank Dan Ksepka and Daniel Thomas, for providing the fossil calibration points and justifications. Together, Alan, Dan and Daniel developed the majority of the taxonomic descriptions in this thesis, and helped make my Ph.D. journey so enjoyable.

I thank Nic Dussex, Ludo Dutoit and Alana Alexander who helped run the demographic analyses. I extend special gratitude to Nic, Ludo and Alana, who assisted with many more analyses than they had signed up for, and whose contributions were key to several manuscripts.

I am grateful to the collaborative nature of Guojie Zhang and Hailin Pan. Guojie and Hailin provided access to several new penguin mitochondrial genomes, without which my project would have been far less comprehensive.

I thank Coté Frugone, who is currently undertaking a Ph.D. project on crested penguin genetics herself. I am glad that we collaborated rather than competed, and for me, this motivated me to extend my project beyond crested penguins.

The samples used in my thesis originate from some of the most remote and inaccessible locations on our planet, with considerable funding having been sunk into their collection over many years. These samples are therefore irreplaceable and invaluable to our research community and I am incredibly grateful that they were offered to me for my thesis. For sample collection I thank Catherine Bazjak, Andy Black, Jenny Booth, Yves Cherel, Anne-Sophie Coquel, Richard Cuthbert, Nina Dehnhard, Ursula Ellenberg, Graeme Elliott, Mike Fawcett, Steven Fiddaman, Sarah Fraser, Tom Hart, Johanna Hiscock, Dave Houston, Holly Irvine, Alastair Judkins, Hotte Mattern, Thomas Mattern, Gary Miller, Colin Miskelly, Kyle Morrison, Marion Nicolaus, Paul Nolan, Sam Petersen, Michael Polito, Petra Quillfeldt, Chris Rickard, Peter Ryan, Paul Sagar, Paul Scofield, Adrian Smith, Alan Tennyson, David Thompson, Kath Walker and Barbara Wienecke, and for access to samples I thank Juan Bouzat, Emma Burns, Alan Cooper, Vanesa de Pietri, Andrea Faris, Brian Gill, Leo Joseph, Pierre Jouventin, Kieren Mitchell, Robert Palmer, Hailin Pan, Nic Rawlence, Matt Rayner, Paul Scofield, Lara Shepherd, Alan Tennyson, Trudi Webster, Jamie Wood, Guojie Zhang, the Museum of New Zealand Te Papa Tongarewa, Canterbury Museum, Auckland War Memorial Museum and Otago Museum. I thank all of you for trusting me to use these precious samples to their full potential.

For laboratory support, I thank Nic Rawlence at the University of Otago, Karen Boot, Nic Bolstridge, Carina Davis, Daggi Goeke, Caroline Mitchell, Duckchul Park, Ana Podolyan, Kat Trought and Jamie Wood at Manaaki Whenua Landcare Research, and Pere Bover, Alan

Cooper and Kieren Mitchell at the University of Adelaide. I am grateful to Andrzej Kilian for providing advice on how to prepare modern samples for downstream sequencing.

Analytical support was provided by Alana Alexander, Gemma Clucas, Nic Dussex, Ludo Dutoit, Arthur Georges, Bernd Gruber, Andrzej Kilian, Michael Knapp, Dan Ksepka, Kieren Mitchell, Hailin Pan, Nic Rawlence, Ben Roberts, Albert Savary, Anthony Shaw, Daniel Thomas, Jonathon Ung, Jamie Wood, Xander Xue and Jane Younger. I utilised the CIPRES Scientific Gateway and the New Zealand eScience Infrastructure for analyses.

Written work benefited from the feedback by all co-authors and reviewers. In particular, I thank Nic Dussex, Ludo Dutoit, Alana Alexander, Dan Ksepka, Kieren Mitchell, Alan Tennyson, Jon Waters, Janet Wilmshurst and Jamie Wood, and my mum - Inge Kowanko.

My project benefited from discussions with my supervisors, my co-authors, my advisors - Michael Knapp and Nic Rawlence, and Dave Ainley, Isabelle Ansorge, Luciano Beheregaray, Alex Boast, Chris Brauer, Frank Cross, Duncan Cunningham, Lea de Nascimento, Vanesa de Pietri, Ewan Fordyce, Crid Fraser, Alicia Grealy, Stef Großer, Mike Hickerson, Kealoha Kinney, Daphne Lee, Bill Lee, Bastien Llamas, Graeme Loh, Robin Long, Matt McGlone, Bruce McKinlay, Edward Mitchell, Colin Meurk, Nick Mortimer, Sandra Nogué, Robert Poulin, Kalinka Rexer-Huber, Marcus Richards, Sarah Richardson, Jessica Rivera-Perez, Bruce Robertson, Daniel Ruzzante, Anna Santure, James Scott, Neil Silverwood, Antje Steinfurth, Ian Smith, Paul Sunnucks, Ayla van Loenen, Trudi Webster, Lauren White, David Whitehead, Otto Whitehead, Kerry-Jane Wilson, Eric Woehler and Trevor Worthy.

I am grateful to Jean-Claude Stahl and Jamie Wood for providing photographs of specimens, and to Sean Murtha for the incredible painting of our two newly described penguin taxa.

I was based between the University of Otago's Department of Zoology and Manaaki Whenua Landcare Research's Long Term Ecology Lab. I thank Sue Scheele, Janet Wilmshurst and Jon Waters for arranging this affiliation, and the Ecosystem Resilience RPA MBIE SSIF funding for facility support. Important aspects of my project only ran smoothly thanks to Robyn Cameron, Daile Hendry, Vivienne McNaughton, Tania King and Nic Rawlence, who processed my orders, arranged reagents and organised postage. I am also lucky to have a fantastic network of friends who welcomed me into their homes when I visited for sampling, discussions or conferences. I thank Angela Chen and Josh Marks, Anthony and Julia Eyles, Chris Garden and Jen Lawn, Kat Johnson and Tristan Marks, Daphne and Bill Lee, Alice Liddell, Graeme Loh and Sue Maturin, Thomas Mattern and Ursula Ellenberg, Nic Rawlence and Maria Zammit, Emilie and Michael Robinson and Trudi Webster and Will Rayment for your hospitality.

My research was supported by the University of Otago, the University of Oxford, the University of Copenhagen, Manaaki Whenua Landcare Research, BGI and the National Museum of New Zealand Te Papa Tongarewa. My co-authors were supported by the Marsden Fund, the Australian Research Council, the United States National Science Foundation, a Rutherford Discovery Fellowship, Quark Expeditions, Southern Discoveries, Cheesemans Ecology Safaris, Golden Fleece Expeditions, New Island Conservation Trust, Deutsche Forschungsgemeinschaft, The Citadel Foundation, The Dalio Foundation, donations to Penguin Watch and the Institut Polaire Français Paul Emile Victor. I received funding from The Royal Society of New Zealand Hutton Fund, The Ornithological Society of New Zealand and an Alumni of Otago in America Award. I was supported by an Australasian Seabird Group grant, an Otago University Postgraduate Award, an Otago University Postgraduate Publishing Bursary and contracts from Manaaki Whenua Landcare Research and Flinders University. I thank Janet Wilmshurst, Phil Lyver, Dean Anderson and Narelle Hunter for these contracts.

I was able to attend conferences through the Otago Ecology Fund, the Otago Community Trust Grant and a University of Otago Postgraduate Conference Grant. I presented preliminary results at the 2016 and 2018 *Oamaru Penguin Conference*, the 2017 and 2019 *Genetics Society of Australasia*, the 2017 *Conference of Australasian Vertebrate Evolution, Palaeontology and Systematics* and the 2019 *International Penguin Conference*. Although I enjoyed being on the *Conference of Australasian Vertebrate Evolution, Palaeontology and Systematics* organising committee, I do not recommend organising a conference during a Ph.D.

My research was approved by the Otago University Animal Ethics Committee 61/2016, the Oxford University Local Animal Ethics Committee, the University of Western Australia Animal Ethics Committee, the Woods Hole Oceanographic Institution Animal Care and Use Committee, the IPEV ethics committee, BGI Ethics Committee and the Zoological Society of London. Samples were collected under New Zealand Department of Conservation Permits (OT-25557-DOA, IACUC-18958.00, 32202-FAU, 35682-FAU, 37312-LND, 50437-DOA, 50436-FAU, 50464-DOA, and 38882-RES), New Zealand Ministry of Primary Industries (2015056535, 2016060908, 2017064905), a Permit to Possess Threatened Fauna for Scientific Purposes No. TFA 15086, a DPIWE Permit to Take Wildlife for Scientific Purposes No. FA05246, the Tristan Da Cunha Conservation Department, the South African Department of Environmental Affairs, the Government of South Georgia and the South Sandwich Islands Restricted Activity Permits (GSGSSI RAP 2018/018, GSGSSI), the United States National Science Foundation Department of Polar Programs ACA permits (ACA 2016-011, ACA 2016-012), the Falkland Islands Government (R05/2009, R014/2006) and United Kingdom Antarctic

Permits. I thank my co-authors for organising the vast majority of these permits, and Neil Fowke and Bruce McKinlay for New Zealand permit approval. Consultation was undertaken with the Ngāi Tahu Research Consultation Committee.

My Ph.D. has opened new doors and enabled me to visit many interesting and remote places. Thomas Mattern and Ursula Ellenberg welcomed me on their 2016 fieldwork to Milford Sound and Jackson Head on New Zealand's West Coast, and it was on this trip that I saw Fiordland crested penguins for the very first time. I was awarded a Heritage Fund in 2016, which took me to The Snares, Campbell and Auckland Islands on the *Spirit of Enderby*, where I saw yellow-eyed and Snares crested penguins for the very first time. I participated in the National Science Foundation Antarctic Biology Training Programme at McMurdo Base in 2018, with very clever early-career scientists, and I saw emperor and Adélie penguins for the very first time. I returned to Antarctica in 2019 with Jamie Wood, Morgan Coleman and Chris Long, where we travelled through Scott Base and Mario Zucchelli Station, sampling ancient penguin rookeries at Terra Nova Bay, Cape Hallett and Cape Adare, with plans to return to Ross Island in early 2020.

I am grateful to Guojie Zhang, Phil Lyver and Dean Anderson, who after only two years into my Ph.D. journey, shone the light at the end of the tunnel, with very cool postdoctoral projects to dive into. Without these incentives, I may still be beavering away at this thesis.

I thank my friends, Petra Quillfeldt, Juan Masello and Yoshan Moodley, who led the research for my first ever co-authored publication, which explored genomic evolution of sub-Antarctic prions. Being involved in that project was an invaluable starting point for me. It provided me with crucial contacts to form new collaborations during my Ph.D., and I don't think I would have had access to all of the penguin samples without that stepping stone.

I am extremely grateful that our family and friends have visited us in New Zealand. Your company has been invaluable to my productivity and has kept me grounded. In particular, I thank my mum and dad - Inge Kowanko and Steve Cole, my Oma - Rosemarie Kowanko, my cousins - Yve Cole, Jon, Mika and Saskia Herd, my new family - Loretta, Russell, Daniel and Steph Wood and Miguel Tapia for your support, encouragement and company.

Finally, I extend my deepest gratitude to my best friend, my colleague and my husband, Jamie Wood. At work, Jamie gave me the space and encouragement to be an independent researcher, providing technical, analytical and written advice only when asked. At home, Jamie ensured that I had a balanced life, built me a bouldering wall in our garage, did more than his fair share of domestic chores, and 'engaged' in my favourite topic of conversation: 'penguins'.

Table of Contents

List of Figures

Chapter 5

List of Supplementary Figures

Chapter 4

List of Supplementary Tables

Chapter 2

Chapter 3

Chapter 4

Chapter 5

Abbreviations

A	Adenine
ABC	Approximate Bayesian Computation
ACAD	Australian Centre for Ancient DNA
ACC	Antarctic Circumpolar Current
aDNA	Ancient DNA
AIC	Akaike Information Criteron
AMOVA	Analysis of MOlecular VAriance
ANWC	Australian National Wildlife Collection
APF	Antarctic Polar Front
BCE	Before common era (interchangeable with BC)
BEAST	Bayesian Evolutionary Analysis Sampling Trees
BGI	Beijing Genomics Institute
BIC	Bayesian information criteron
BOLD	Barcode of Life Database
BP	Before present (1950 CE)
bp	Base pairs
BSA	Bovine serum albumin
BSP	Bayesian Skyline Plot
BWA	Burrows-Wheeler Aligner
C	Cytosine
cal	Calibrated
CE	Common era (interchangeable with AD)
cf.	Compared with/to
CI	Confidence interval
CIPRES	Cyber Infrastructure for Phylogenetic RESearch
cm	Centimetre
COI	Cytochrome Oxidase subunit 1 gene
CR	Control Region
Cyt*b*	Cytochrome *b* gene
DAPC	Discriminant analysis of principal components
DArT	Diversity Arrays Technology
DArT-seq™	Diversity Arrays Technology Pty Ltd.
DNA	Deoxyribonucleic acid
dNTPs	Deoxyribonucleotide triphosphate
ESS	Effective sample size
EtOH	Ethanol
e.g.	For example
F_{ST}	Population differentiation (fixation index)
G	Guanine
GBS	Genotyping by Sequencing
G_{IS}	Inbreeding coefficient
GTR	General Time Reversible
h	Hour
H_E	Expected heterozygosity
HiFi	High Fidelity
HKY	Hasegawa-Kishino-Yano
H_O	Observed heterozygosity
HPD	Highest posterior density
i.e.	In other words
K	Genetic clusters
Ka	Thousand years
km	Kilometres
kya	Thousand years ago
LGM	Last Glacial Maximum
M	Molar (Moles/Litre)
m	Metre
Ma	Million years
MAF	Minor allele frequency
MCMC	Markov Chain Monte Carlo
mg	Milligram
$MgSO_4$	Magnesium sulphate
min	Minutes
mitogenome	Mitochondrial genome
ML	Maximum-likelihood
mL	Millilitre
mM	Millimolar
mm	Millimetre
mtDNA	Mitochondrial DNA
MUSCLE	Multiple sequence comparison by log-expectation
MWLR	Manaaki Whenua Landcare Research
mya	Million years ago
n	Number of samples
NA	Not applicable
ND3	NADH-ubiquinone oxidoreductase chain 3
N_e	Effective population size
NeSI	NZ eScience Infrastructure
ng	Nanogram
NGS	Next Generation Sequencing
nM	Nanomolar
NMNZ	National Museum of New Zealand Te Papa Tongarewa
nov.	New species
NUMTs	Nuclear mitochondrial DNA segments
NZ	New Zealand
PBS	Phosphate buffered saline
PC	Principal component
PCoA	Principal components analysis
PCR	Polymerase chain reaction
ProtK	Proteinase K
QL buffer	Queen's lysis buffer
RAD-seq	Restricted site Associated DNA Sequencing
RAG1	Recombination Activating 1 gene
RAxML	Randomized Axelerated Maximum Likelihood
RNase	Ribonuclease
RO H_2O	Reverse osmosis water
RSA	Rabbit serum albumin
rtPCR	Real-time PCR
s	Second
SE	Standard error
SFS	Site frequency spectrum
SNAPP	SNP and AFLP Package for Phylogenetic analysis
SNP	Single nucleotide polymorphism
sp.	Species
ssp.	Subspecies
StAMPP	Statistical Analysis of Mixed Ploidy Populations
STF	Subtropical Front
T	Thymine
T	The time in the past when N_e changed
TBE	Tris/Borate/Ethylenediaminetetraacetic acid buffer
U	Units
µL	Microlitre
µM	Micromolar
VCF	Variant call format
X	Times (reagent concentration)
12S	12S ribosomal RNA
^{14}C	Radiocarbon years
16S	16S ribosomal RNA
3'	3 prime end
5'	5 prime end
°C	Degrees Celcius
°S	Degrees South

Nomenclature

Aptenodytes forsteri	Emperor penguin
Aptenodytes patagonicus	King penguin
Pygoscelis adeliae	Adélie penguin
Pygoscelis antarctica	Chinstrap penguin
Pygocelis papua	Gentoo penguin
Eudyptula minor	New Zealand little blue penguin
Eudyptula novaehollandiae	Australian fairy penguin
Spheniscus mendiculus	Galápagos penguin
Spheniscus magellanicus	Magellanic penguin
Spheniscus humboldti	Humboldt penguin
Spheniscus demersus	African penguin
Megadyptes antipodes antipodes	Yellow-eyed penguin
Megadyptes antipodes richdalei	Chatham Island yellow-eyed penguin (extinct)
Megadyptes antipodes waitaha	Waitaha penguin (extinct)
Eudyptes chrysolophus chrysolophus	Macaroni penguin
Eudyptes chrysolophus schlegeli	Royal penguin
Eudyptes chrysocome	Western rockhopper penguin
Eudyptes filholi	Eastern rockhopper penguin
Eudyptes moseleyi	Northern rockhopper penguin
Eudyptes pachyrhynchus	Fiordland-crested penguin
Eudyptes robustus	Snares-crested penguin
Eudyptes sclateri	Erect-crested penguin
Eudyptes warhami	Chatham Island crested penguin (extinct)

Chapter 1

Background and aims of this thesis

General Introduction on island evolution and biogeography

Oceanic islands represent natural laboratories for the study of evolutionary and biogeographic processes (Darwin, 1859; MacArthur & Wilson, 1967; Whittaker & Fernández-Palacios, 2007). For instance, the tectonic or volcanic emergence of young islands and archipelagos can provide new unoccupied niches for dispersing species to colonise, leading to subsequent adaptation and diversification (Fleischer *et al*., 1998; Gathorne-Hardy & Jones, 2000; Mendelson & Shaw, 2005; Gillespie *et al*., 2012). The establishment of just a few founder individuals may ultimately facilitate the evolution of new taxa (Waters *et al*., 2013). Island biological assemblages are often characterised by spectacular adaptive radiations (Losos *et al*., 1998; Grant & Grant, 2002; Gillespie, 2004; Lerner *et al*., 2011; see *Figure 1*), high levels of endemism (Kier *et al*., 2009), loss of flight (McNabb, 1994), gigantism and dwarfism (Lomolino, 2005; Knapp *et al*., 2019). Isolated island populations are often descended from small founding populations, and can also experience rapid genetic drift, leading to relatively low levels of genetic diversity (Clegg *et al*., 2002). Oceanic island assemblages are therefore vastly different to those on the continents (Stuart *et al*., 2012). Additionally, while island ecosystems comprise almost a third of the world's biodiversity hotspots (Myers *et al*., 2000), many of these systems have been highly impacted by human activities, including hunting, habitat destruction, the introduction of predatory species and climate change (Kier *et al*., 2009; Nogué *et al*., 2017; Valente *et al*., 2019).

The vast majority of island biology research has taken place on single or localised island groups, with classic examples including the Galápagos (Darwin, 1859; Grant & Grant, 2002; Edgar *et al*., 2004; Parent *et al*., 2008; Benavides *et al*., 2009), Hawaiian (Wilson, 1963; Loope *et al*., 1988; Gillespie, 2004; Cowie & Holland, 2006; Lerner *et al*., 2011; see *Figure 1*) and Macronesian (Juan *et al*., 2000; Emerson, 2002; Rutschmann *et al*., 2017) archipelagos. While many evolutionary and biogeographic processes acting on local island assemblages have been well-documented (Shaw & Gillespie, 2016), it remains unclear how such processes might scale up across broader spatial scales. For example, many islands in the vast Southern Ocean are separated by thousands of kilometres of open ocean, yet remain highly connected by strong winds and ocean currents, sharing cool, wet and windy climates (Bergstrom & Chown, 1999),

and have been suggested to represent a single circumpolar 'Insulantarctica' biogeographic province (Udvardy, 1987). Many Southern Ocean islands are geologically young (less than five million years old [Ma]; Hodgson *et al*., 2014; Moon *et al*., 2017). At least for highly dispersive species, it is conceivable that the evolutionary and biogeographic processes often demonstrated across local island archipelagos (e.g. founder speciation; Waters *et al*., 2013) may also be important across the Southern Ocean. However, whether this is true remains unclear. In fact, there remains a degree of uncertainty about the biogeographic origins of many Southern Ocean endemic terrestrial species (Moon *et al*., 2017). For example, molecular clock studies have suggested divergence dates for some species that are older than the age of islands on which they currently live (Heenan *et al*., 2010). Further studies are therefore needed to resolve general patterns about how terrestrial species disperse and evolve throughout the Southern Ocean.

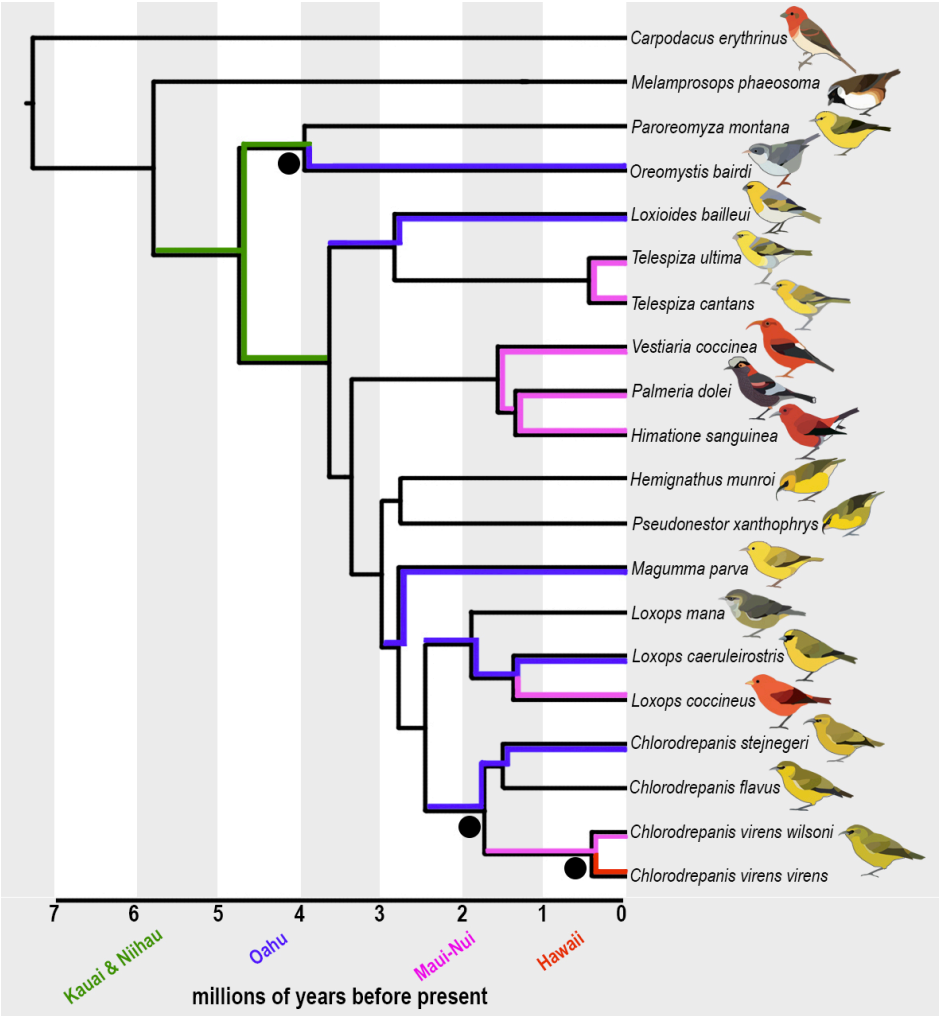


Figure 1. Island formation and isolation plays a key role in biological diversification. For instance, the adaptive radiation of Hawaiian honeycreepers evolved following the emergence of Kauai Island and Niihau Island, with subsequent diversification events occurring after Oahu Island, Maui-Nui Island and Hawaii Island emerged. Each species evolved a distinctive beak morphology to exploit specific fruits, insects and nectar present on particular islands. The phylogenetic tree is adapted from Lerner *et al.* (2011) and is based on mitogenomes. The three calibration points are based on the geological age of islands, and are marked with a black ellipse. Island ages are marked below the scale bar. Note *Chlorodrepanis virens* encompasses two subspecies, but only one illustration is shown.

The Southern Ocean

While the high-latitude Northern Hemisphere Arctic Ocean is largely surrounded by land, the vast (and relatively unexplored) Southern Ocean completely rings the entire Antarctic continent (Fraser *et al*., 2011; see *Figure 2*). The distinctive circumpolar flow of the Southern Ocean formed approximately 30 million years ago (mya) when Antarctica and South America separated, leading to the formation of Drake Passage and the world's strongest surface current, the Antarctic Circumpolar Current (ACC; Barker & Burrell, 1977). Unimpeded by any major land mass, the Southern Ocean is characterised by persistent westerly (eastward) winds and currents. While the Antarctic Polar Front (APF) encircles Antarctica, cold northward-flowing Antarctic waters mix and sink beneath the relatively warmer sub-Antarctic waters, driving upwelling (Armour *et al*., 2016). Upwelling creates zones high in nutrients, supporting abundant marine life (Holm-Hansen, 1985), including phytoplankton, krill, squid, fish, whales, seals, seabirds and penguins. The importance of the Southern Ocean therefore has a critical influence on global climate and biodiversity, due to its storage and distribution of heat, carbon, oxygen and nutrients, coupled with its connection to the Pacific, Indian and Atlantic Oceans (Marshall & Speer, 2012; see *Figure 2*).

While the exact definition of the Southern Ocean remains controversial, it is herein used broadly and inclusively to refer to the entire body of water surrounding the Antarctic continent (*Figure 2*), including the extensive current systems that bring cold water northward to the coastal environments of southern Africa, southern Australia, southern New Zealand and southern South America. Dotted throughout the vast Southern Ocean are dozens of volcanic and tectonic island archipelagos (*Figure 2*). While these islands are geographically isolated from the southern continents of Australia, Africa, South America and Antarctica, they host abundant endemic biota (Chown *et al*., 1998) and represent important nesting grounds for a large proportion of the world's seabird diversity (Bost *et al*., 2009; Woehler *et al*., 2011), including albatrosses, petrels, skuas, gulls, terns and penguins (Onley & Scofield, 2008; see *Figure 2*). With contrasting geological and climate histories, combined with recent anthropogenic impacts, these Southern Ocean islands are natural laboratories for studying evolutionary and biogeographic processes across a vast circumpolar scale.

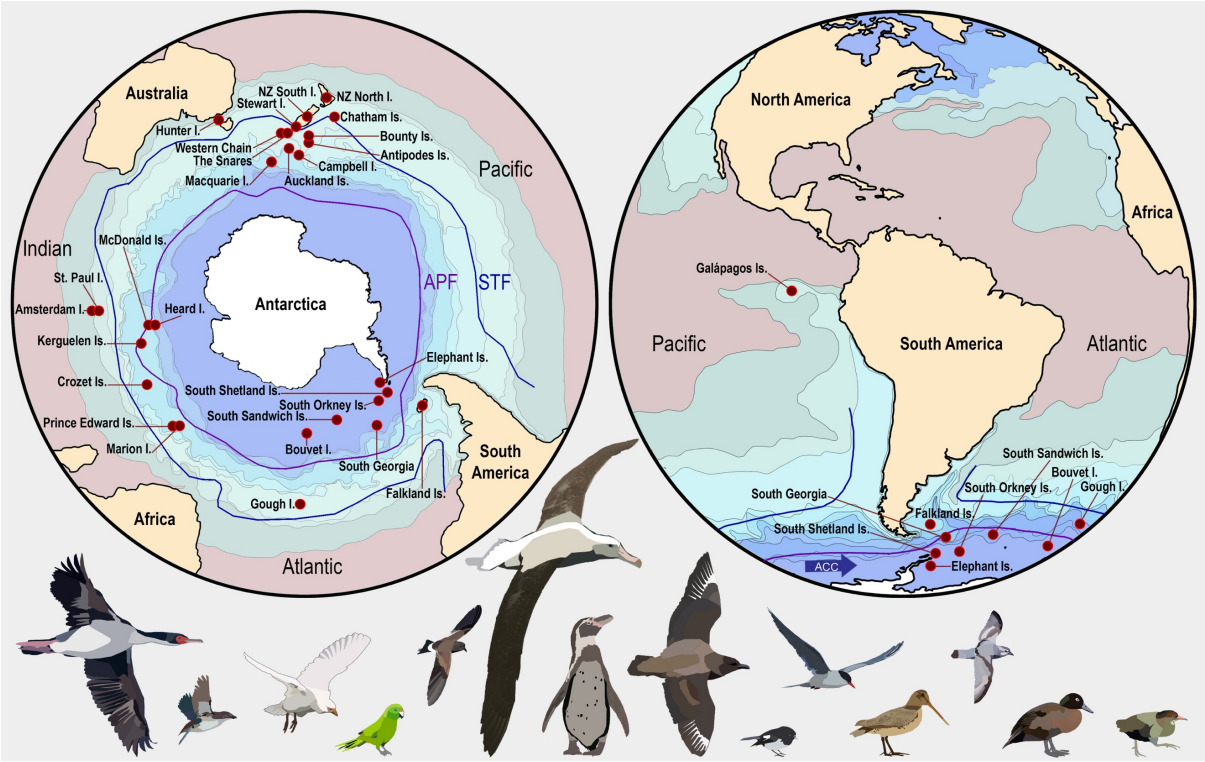


Figure 2. The Southern Ocean is critical for the world's biodiversity and climate. The maps are adapted from Earth (https://earth.nullschool.net) and are based on average sea surface temperatures during the Southern Hemisphere winter equinox (June 22, 2018). Water south of the Antarctic Polar Front (APF; lilac) is approximately -2°C; water south of the Subtropical Front (STF; aqua) is approximately 8°C, and water north of the STF (blue) is approximately 20°C. Tropical water (pink) is approximately 28°C. The eastwards Antarctic Circumpolar Current (ACC) is shown. Southern Ocean islands are marked. Some representative species from 13 avian families that inhabit islands in the Southern Ocean are illustrated, from left to right; Bounty Islands shag (*Phalacrocorax ranfurlyi*; Phalacrocoracidae), common diving petrel (*Pelecanoides urinatrix*; Procellariidae), Antarctic snowy sheathbill (*Chionis albus*; Chionidae), Antipodes Islands parakeet (*Cyanoramphus unicolor*; Psittacidae), Wilson's storm petrel (*Oceanites oceanicus*; Oceanitidae), wandering albatross (*Diomedea exulans*; Diomedeidae), Humboldt penguin (*Spheniscus humboldti*; Spheniscidae), Antarctic skua (*Catharacta antarctica*; Stercorariidae), Auckland Islands tomtit (*Petroica macrocephala marrineri*; Petroicidae), Antarctic tern (*Sterna vittata*; Laridae), Auckland Islands snipe (*Coenocorypha aucklandica aucklandica*; Scolopacidae), fairy prion (*Pachyptila turtur*; Procellariidae), Auckland Islands teal (*Anas aucklandica*; Anatidae) and Auckland Islands rail (*Lewinia muelleri*; Rallidae).

Biological colonisation and diversification on Southern Ocean islands

In contrast to the Southern Hemisphere's ancient 'Gondwanan' continents, many of the Southern Ocean islands have relatively recent geological histories, emerging only during the last five million years or so, as a result of tectonic or volcanic activity (McDougall & Ollier, 1982; Maund *et al.*, 1988; Paulay, 1994; Adamson *et al.*, 1996; Campbell *et al.*, 2008; Scott & Turnbull, 2019). Soon after emergence, isolated islands can be rapidly colonised by diverse dispersing taxa, which together form new ecosystems, creating new opportunities for local adaptation and evolutionary diversification (Fleischer *et al.*, 1998; Garthorne-Hardy & Jones, 2000; Mendelson & Shaw, 2005; Gillsepie *et al.*, 2012). For example, the young 300 thousand year old (Ka) New Zealand Antipodes Islands (Scott & Turnbull, 2019) provides nesting

grounds for several locally endemic species, including the Antipodes Islands parakeet (*Cyanoramphus unicolor*), Reischek's parakeet (*C. hochstetteri*), the largely flightless Antipodes Island snipe (*Coenocorypha aucklandica meinertzhagenae*) and the majority of the nesting grounds for the erect-crested penguin (*Eudyptes sclateri*). Similarly, the young (three million year old) New Zealand Chatham Islands provide the only nesting grounds for many bird species, including the black robin (*Petroica traversi*), Forbes' parakeet (*Cyanoramphus forbesi*), Chatham snipe (*Coenocorypha pusilla*) and Chatham albatross (*Thalassarche eremita*), and until recently was home to at least 22 taxa that are now extinct (Tennyson & Millener, 1994). While past studies on endemic Chatham Island plants (Heenan *et al*., 2010), and the extinct Chatham duck (*Anas chathamica*; Mitchell *et al*., 2014a) and Chatham kaka (*Nestor chathamensis*; Wood *et al*., 2014) have hinted that the uplift of the Chatham Islands may have potentially played a role in species diversification, it remains conceivable that recent, circumpolar island emergence has played a key role in the evolutionary diversification of other endemic Southern Ocean species. These Southern Ocean situations potentially mirror well-studied examples of island diversification in northern latititudes, such as the Galápagos and Hawaiian archipelagos (see *Figure 1*).

While many Southern Ocean islands are young, all were dramatically impacted by Quaternary inter-glacial and post-glacial climate fluctuations during the last two and a half million years (Hodgson *et al*., 2014), resulting in sea-level changes and sometimes glaciation, especially during the Last Glacial Maximum (LGM; 18 – 25 thousand years ago [kya]; Jouzel *et al*., 2007; Hodgson *et al*., 2014). The biological impacts of such climatic changes can be substantial. For example, numerous studies have suggested that during inter-glacial periods, many cold-adapted species retreat to mid-latitude refugia, and subsequently expand towards the poles during inter-glacial periods (Fraser *et al*., 2011). Such shifts may lead to dramatic changes in both population size and biogeographic range (Cristofari *et al*., 2018). While post-LGM population expansions have been extensively studied in Northern Hemisphere biota (Parmesan & Yohe, 2003), they remain relatively poorly understood in Southern Ocean ecosystems. The LGM saw extensive glacial and sea-ice south of the Southern Hemisphere's Subtropical Front (STF; Gersonde *et al*., 2005; Esper & Gersonde, 2014), presumably leading to the extirpation of many Antarctic and sub-Antarctic terrestrial and marine populations, as coastal habitat was replaced by ice, and previously abundant food sources disappeared (de Bruyn *et al*., 2009; Fraser *et al*., 2011). Species with high dispersal ability presumably avoided extinction by moving to refugial regions in higher altitudes, or by moving away from the polar regions to mid-latitude refugia (Hewitt, 2004; Fraser *et al*., 2009; Fraser *et al*., 2012; Younger *et al*., 2015a; Younger *et al*., 2015b), while some species with cold thermal tolerance may have survived very locally south of the

STF in isolated ecosystems hosted by polynyas (open water surrounded by sea ice), e.g. king (*Aptenodytes patagonicus*) and Adélie penguins (*Pygoscelis adeliae*) in the Ross sea (Brambati *et al*., 2002; Thatje *et al*., 2008; Younger *et al*., 2015c; Clucas *et al*., 2016; Emslie *et al*., 2018). While polar regions became heavily glaciated, landmasses further north experienced lower sea levels by approximately 120 – 135 metres (Clark & Mix, 2002), which may have created new opportunities for the establishment of northward-dispersing refugia, which were then able to persist locally. Following the LGM, global climate warmed once again, high-latutude Southern Ocean islands became de-glaciated, and sea-levels rose. Those high-latitude Southern Ocean islands became habitable once again and refugial populations of cold-adapted species moved back towards the southern regions (Fraser *et al*., 2011). It is therefore conceivable that the widespread distributions of many Southern Ocean species may have been promoted by relatively recent post-LGM warming, facilitating circumpolar dispersal into southern regions freed from ice (Nikula *et al*., 2010; see also *Figure 3*). For example, multiple islands and coastlines south of the STF provide breeding grounds for widespread albatrosses (*Diomedea*, *Thalassarche*, *Phoebetria* genera), petrels (*Macronectes*, *Thalassoica*, *Daption*, *Pagodroma*, *Procellaria* genera), fulmars (*Fulmarus*), prions (*Pachyptila*) and penguins (*Eudyptes*, *Pygoscelis* and *Aptenodytes* genera). Nevertheless, the few studies that have directly explored post-glacial population expansions in Southern Ocean fauna have revealed contrasting patterns (de Bruyn *et al*., 2009; Li *et al*., 2014; Trucchi *et al*., 2014; Younger *et al*., 2015a; Younger *et al*., 2015c; Younger *et al*., 2016; Corrigan *et al*., 2016; Cristofari *et al*., 2016; Cristofari *et al*., 2018; Frugone *et al*., 2018; Carrea *et al*., 2019; Chau *et al*., 2019).

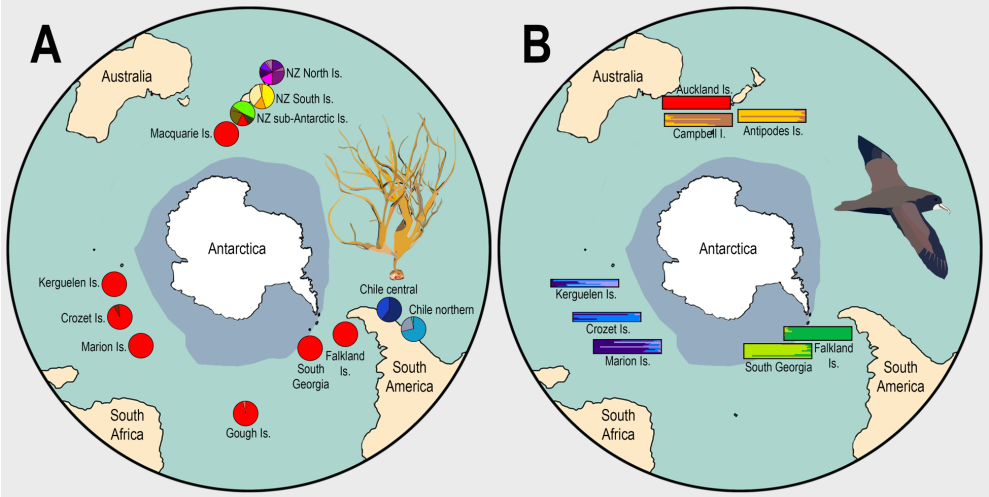


Figure 3. Southern Ocean species exhibit contrasting patterns of genetic structure. **A)** has been adapted from Fraser *et al*. (2009), and shows southern bull-kelp (*Durvillaea antarctica*) haplotypes based on COI. Islands south of the Subtropical Front (STF) are genetically homogeneous (shown in red). **B)** is adapted from Rexer-Huber (2017), and shows white-chinned petrel (*Procellaria aequinoctialis*) structure plots based on SNPs. Three evolutionary significant units are visible (shown in red/orange, blue and green) and correspond to different geographical regions. The grey surrounding Antarctica shows contemporary summer sea-ice.

Until only the last few hundred years, most isolated Southern Ocean island ecosystems have remained free from the impacts of humans. New Zealand (NZ) provides an excellent case study of the impacts of human arrival on island ecosystems (Waters *et al*., 2017; Valente *et al*., 2019). Human arrival to temperate NZ over the last 780 years (Wilmshurst *et al*., 2008; Wilmshurst *et al*., 2011) caused considerable ecological transformations (Seersholm *et al*., 2018). In temperate NZ and the nearby Chatham Islands, for example, more than 61 avian extinctions have been documented (Tennyson & Martinson, 2007; Boessenkool *et al*., 2008; Wood *et al*., 2014), with evidence for rapid extinction and re-colonisation (see *Figure 4*; Trewick & Worthy, 2001; Boessenkool *et al*., 2008; Collins *et al*., 2014a; Rawlence *et al*., 2015a; Grosser *et al*., 2016; Rawlence *et al*., 2017a), together with range reductions detected in persisting species; e.g. kea (*Nestor notabilis*; Dussex *et al*., 2015), NZ sea lions (*Phocarctos*; Rawlence *et al*., 2016), fur seals (*Arctocephalus*; Salis *et al*., 2016), brown teals (*Anas chlorotis*; Cole & Wood, 2017a) and shags (*Leucocarbo*; Rawlence *et al*., 2017b).

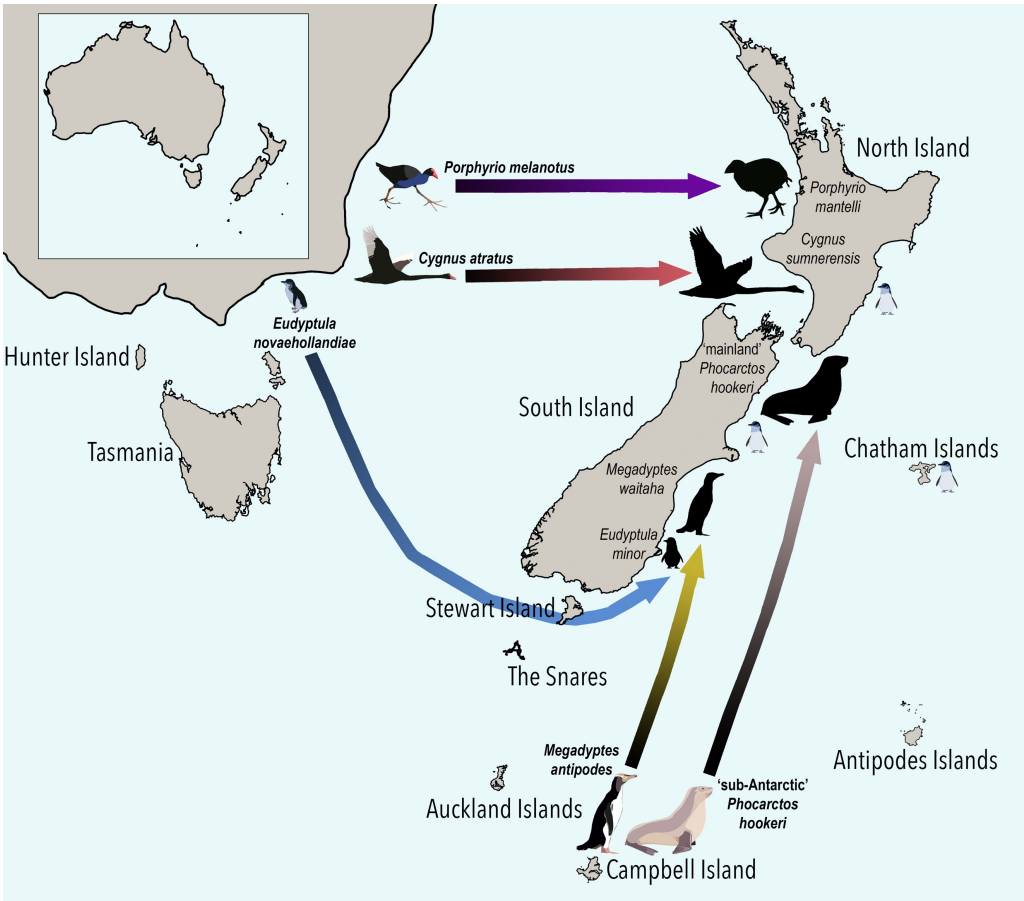


Figure 4. Rapid extinction and recolonisation events in NZ have been uncovered in high dispersal taxa. Examples include yellow-eyed penguins (*Megadyptes*; Boessenkool *et al*., 2008; Rawlence *et al*., 2015a) and NZ sea lions (*Phocarctos*; Collins *et al*., 2014a), which recolonised the NZ South Island from sub-Antarctic lineages, and little penguins (*Eudyptula*; Grosser *et al*., 2016), black swans (*Cygnus*; Rawlence *et al*., 2017a), and swamphens (*Porphyrio*; Trewick & Worthy, 2001) which recolonised NZ from Australian lineages. Extinct taxa or lineages are shown in black. Note the NZ little-blue penguin (*Eudyptula minor*) still occurs throughout much of NZ, but the Otago coastline was replaced by the Australian fairy penguin (*E. novaehollandiae*) following local extinction of *E. minor*.

Threats to Southern Ocean ecosystems

Since their discovery by humans, many Southern Ocean islands have lost substantial biodiversity (Worthy & Holdaway, 2002; Frenot *et al*., 2004; Tennyson & Martinson, 2007; Seersholm *et al*., 2018). In the 1800s, seal, whale and penguin populations on many sub-Antarctic islands were heavily exploited (Falla, 1962; Convey & Lebouvier, 2009; Hofman, 2017), particularly those with permanent settlements, e.g. Kerguelen, Amsterdam Island, Tristan da Cunha, Macquarie Island and Falkland Islands (Armstrong, 1994; Convey & Lebouvier, 2009). Farming of cattle (*Bos taurus*), sheep (*Ovis aries*) and goats (*Capra* spp.) has also caused significant damage to Southern Ocean island ecosystems.

All Southern Ocean islands have some degree of national legal protection (de Villiers *et al*., 2006). However, commercial fishing, the introduction of plants, insects, birds and mammalian predators, e.g. stoats (*Mustela erminea*), cats (*Felis cattus*) and rats (*Rattus* spp.), habitat fragmentation, pollution and global climate change continue to threaten many Southern Ocean island ecosystems (Bergstrom & Chown, 1999; Frenot *et al*., 2001; Gaston *et al*., 2003; de Villiers *et al*., 2006; Angel *et al*., 2008; Forcada & Trathan, 2009; Keogan *et al*., 2018; Mattern & Wilson, 2018). Southern Ocean islands therefore require urgent conservation management strategies.

Biogeography of Southern Ocean penguins

A crucial step towards developing ecosystem-wide conservation priorities in the Southern Ocean is to understand the dynamic histories of endemic biota, in light of major geological, climatic and anthropogenic events. The Southern Ocean hosts numerous ecologically important species, from the keystone Antarctic krill (*Euphausia superba*) to the top marine predators of seals, whales, seabirds, and penguins (Bergstrom & Chown, 1999). Seabirds are sentinels of environmental change in the Southern Ocean (Boersma, 2008), with their health intrinsically linked to the health of marine environments (Burger & Gochfeld, 2004). One way to determine the histories of Southern Ocean taxa is to understand historical and contemporary population connectivity. Past studies have revealed contrasting patterns in population connectivity within widespread seabird species, with white-chinned petrels (*Procellaria aequinoctialis*) and gentoo penguins (*Pygoscelis papua*) clustering into different groups according to their breeding locality (Clucas *et al*., 2014; Levy *et al*., 2016; Vianna *et al*., 2017; Clucas *et al*., 2018; Rexer-Huber *et al*., in press), and other species, such as *Aptenodytes patagonicus*, emperor penguins (*Aptenodytes forsteri*) and macaroni penguins (*Eudyptes chrysolophus chrysolophus*) exhibiting relatively homogeneous population connectivity across their entire distribution

(Clucas *et al*., 2016; Younger *et al*., 2017; Clucas *et al*., 2018; Cristofari *et al*., 2017; Cristofari *et al*., 2018; Frugone *et al*., 2018; Frugone *et al*., 2019).

Although penguins spend much of their lives at sea, most species are tied to natal landmasses to breed. Penguins inhabit every major coastline in the Southern Hemisphere and almost every island archipelago in the Southern Ocean (*Figure 5*), ranging from the diverse environments of the Galápagos archipelago, oceanic-temperate forests of NZ, rocky coastlines of the sub-Antarctic islands and the sea-ice around coastal Antarctica. Approximately 20 extant species are recognised across six well defined genera (*Aptenodytes*, *Pygoscelis*, *Eudyptula*, *Spheniscus*, *Megadyptes* and *Eudyptes*; see *Figures 5 – 6*), with some disagreement among the species boundaries of closely related crested penguin (*Eudyptes*) species (Christidis and Boles, 2008; Frugone *et al*., 2018; Mattern & Wilson, 2018; Frugone *et al*., 2019; Mays *et al*., in press).

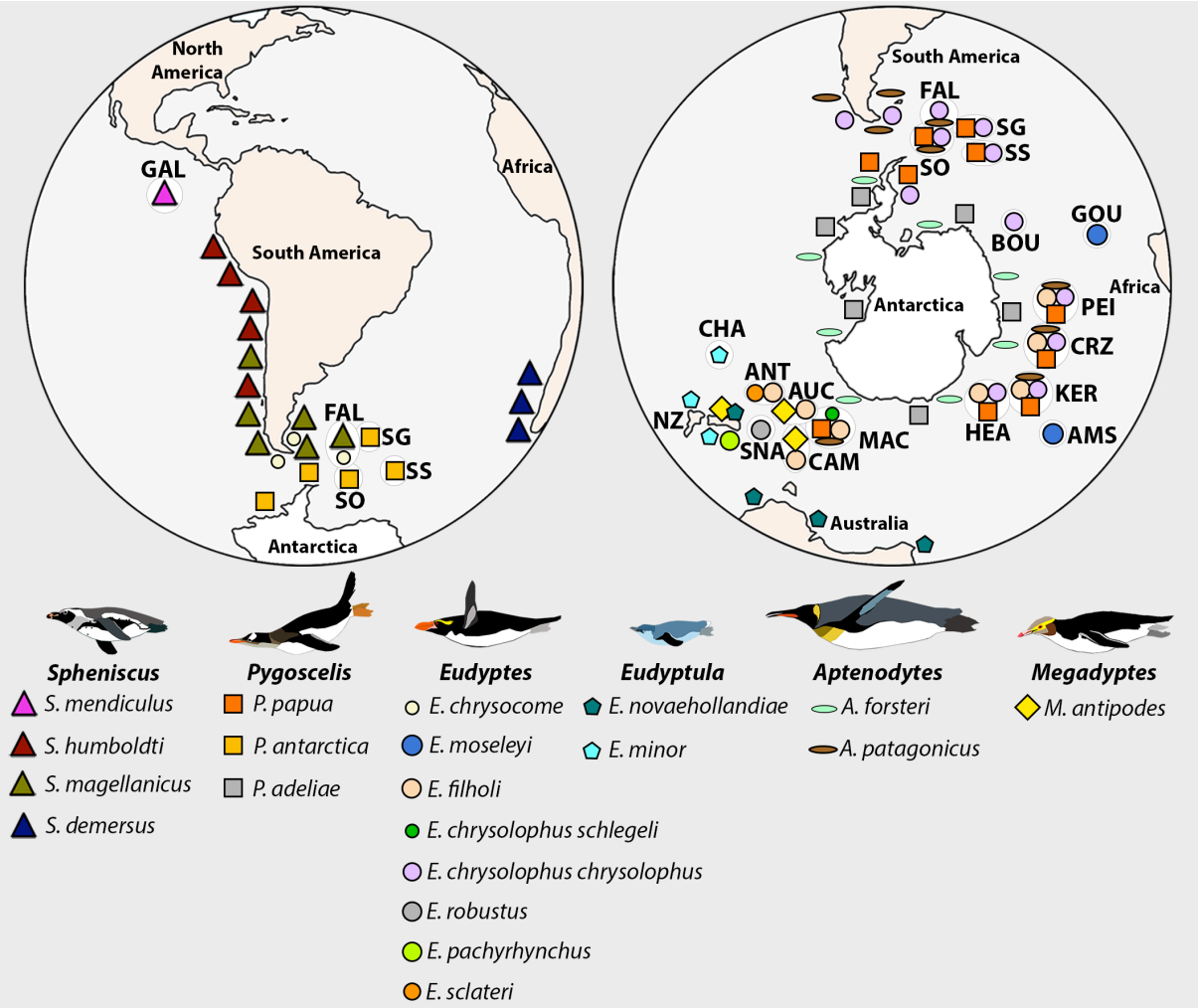


Figure 5. Breeding distributions of extant penguin taxa. Maps are adapted from Ramos *et al*. (2018) and the figure is adapted from Pan *et al*. (in press). GAL: Galápagos Islands; FAL: Falkland Islands; SG: South Georgia; SS: South Sandwich Islands; SO: South Orkney Islands; BOU: Bouvet; GOU: Gough Island; TDC: Tristan da Cunha; PEI: Prince Edward Islands; CRZ: Crozet Islands; KER: Kerguelen Islands; HEA: Heard Island; AMS: Amsterdam and St Paul Islands; MAC: Macquarie Island; CAM: Campbell Island; SNA: Snares; AUC: Auckland Islands; ANT: Antipodes Islands; CHA: Chatham Islands; NZ: New Zealand.

The group has a rich fossil record, extending back more than 60 million years, with more than 50 extinct penguins documented (Slack *et al*., 2006). Many fossil penguins were tall giants (Mayr *et al*., 2017; Richards, 2019), rivalling even the largest extant *Aptenodytes forsteri* (see *Figure 6*). While many penguin taxa are distributed widely across circumpolar sub-Antarctic and Antarctic coastlines, e.g. *Aptenodytes patagonicus*, *A. forsteri*, *Pygoscelis adeliae*, *P. papua*, *Eudyptes chrysolophus chrysolophus* and eastern rockhopper (*E. filholi*), nearly a third of all penguin taxa are endemic to geologically young island archipelagos, e.g. Fiordland-crested (*E. pachyrhynchus*), Snares-crested (*E. robustus*), royal (*E. c. schlegeli*) and Galápagos penguin (*Spheniscus mendiculus*). Penguins therefore, are an ideal group of animals to study how geology, glaciology and human arrival on Southern Ocean islands may have shaped their recent evolution and biogeography.

Past studies of penguins

Over the past two decades, there have been a number of phylogenetic studies aimed at understanding the ancestral relationships among extant and fossil penguins, or to infer the drivers and timing of their evolution (e.g. Baker *et al*., 2006; Gavryushkina *et al*., 2017; Pan *et al*., in press). Past studies mostly used morphological characters, behaviour or short genetic sequences, singularly or in combination. Previous studies suggested that the most recent common ancestor of crown penguins lived approximately 16 – 50 million years ago (Baker *et al*., 2006; Ksepka & Clarke, 2010; Ksepka *et al*., 2012; Subramanian *et al*., 2013). Most recently, Gavryushkina *et al*. (2017) demonstrated that modern penguins may have radiated much more recently than previously estimated, with the basal divergence in the crown clade occurring around 13 million years ago, and most splits leading to extant species occurring only in the past two million years. Previous studies invoked circumpolar ocean currents (Ksepka *et al*., 2006; de Dinechin *et al*., 2009) or Antarctic cooling (Baker *et al*., 2006) as key drivers of evolution, while vicariance (Park *et al*., 2016), ocean currents (Clucas *et al*., 2018; Frugone *et al*., 2018) and glaciation (Clucas *et al*., 2014) are considered drivers of modern biogeographic patterns.

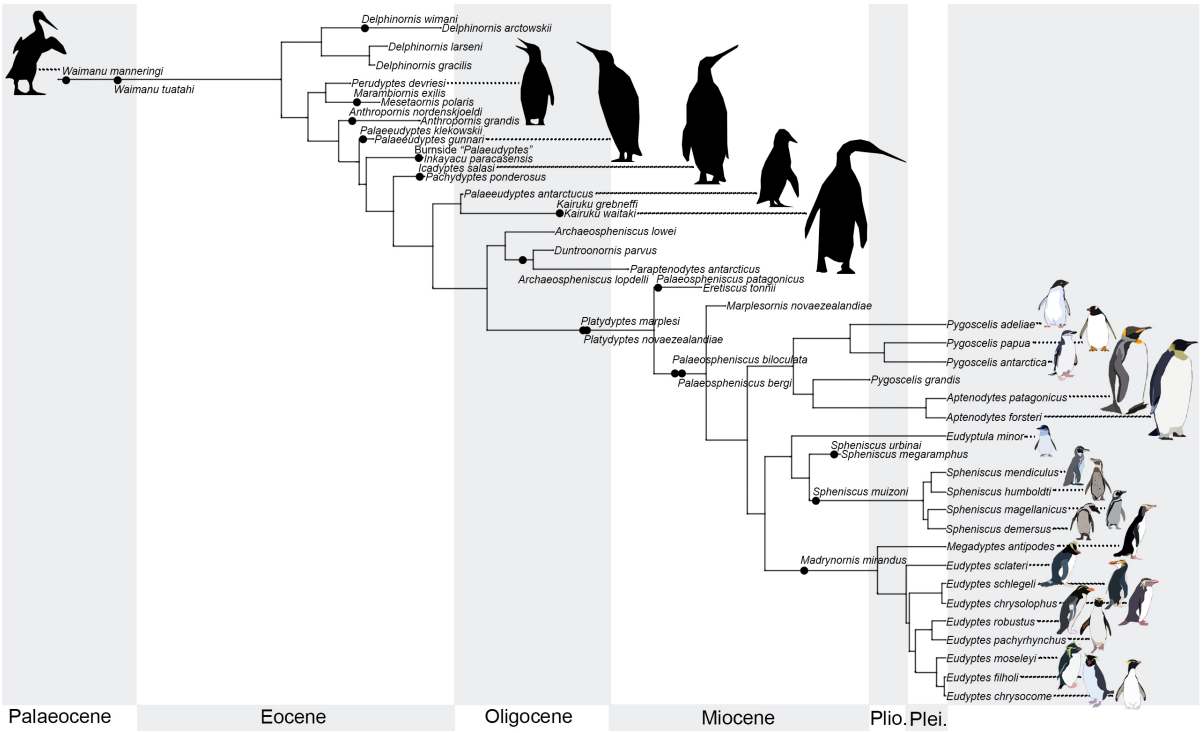


Figure 6. Phylogenetic tree of extant and extinct penguins. The phylogeny (adapted from Gavryushkina *et al*. 2017) is based on the maximum sampled-ancestor clade credibility tree, which uses 202 morphological data characters across extant and fossil taxa, and five mitochondrial and nuclear markers (12S, 16S, COI, cyt*b* and RAG1) for extant taxa. Extant taxa are coloured.

A number of recent studies have used multi-locus genetic or genomic approaches to explore possible relationships within different penguin taxa, or demographic changes over recent Millennia. Most Southern Ocean penguins have now been studied to some level of detail, with the vast majority of taxa exhibiting very limited gene flow across entire distributions; e.g. *Aptenodytes* (Younger *et al*., 2015a; Clucas *et al*., 2016; Cristofari *et al*., 2016; Younger *et al*., 2017; Cristofari *et al*., 2018), *Pygoscelis* (Clucas *et al*., 2014; Pena *et al*., 2014; Levy *et al*., 2016; Vianna *et al*., 2017; Clucas *et al*., 2018; Mura-Jornet *et al*., 2018), *Eudyptes* (Frugone *et al*., 2018; Frugone *et al*., 2019), *Megadyptes* (Boessenkool *et al*., 2008; Boessenkool *et al*., 2009; Boessenkool *et al*., 2010; Lopes & Boessenkool, 2010), *Eudyptula* (Overeem *et al*., 2008; Peucker *et al*., 2009; Burridge *et al*., 2015; Grosser *et al*., 2015; Grosser *et al*., 2016) and *Spheniscus* (Nims *et al*., 2008; Bouzat *et al*., 2009; Schlosser *et al*., 2009; Dantas *et al*., 2018). By contrast, biogeographical patterns in divergent *Pygoscelis papua* populations have been linked to glacial isolation or ocean currents (Clucas *et al*., 2014; Clucas *et al*., 2018), which may have resulted in reduced gene flow and phylogeographical structure.

Penguin systematics and biodiversity change

While several recent studies have examined the evolution and demographic histories of individual penguin taxa (see above), substantial gaps in our knowledge still exist, and no previous study has attempted to provide a major, large-scale synthesis of such histories.

Although penguins are an iconic and relatively well studied seabird assemblage, several new taxa have been recently recognised, including both extant and extinct lineages. For example, phylogenetic analyses of subfossil *Megadyptes* penguins from the NZ South Island revealed a cryptic extinction event, where the Waitaha penguin (*M. waitaha*) was most likely rapidly hunted to extinction, only to be replaced by the extant yellow-eyed penguin (*M. antipodes*) (Boessenkool *et al*., 2008). Similarly, *Eudyptula* penguins also comprise two species; the NZ little blue penguin (*E. minor*) and the Australian fairy penguin (*E. novaehollandiae*) (Grosser *et al*., 2017), while the taxonomic boundaries between extant *Eudyptes* penguins, particularly between *E. chrysolophus chrysolophus*/*E. c. schlegeli*, the western/eastern/northern rockhopper penguins (*E. chrysocome*/*E. filholi*/*E. moseleyi*) and *E. pachyrhynchus*/*E. robustus* remain contested (Christidis & Boles, 2008; Frugone *et al*., 2018; Mattern & Wilson, 2018; Frugone *et al*., 2019; Mays *et al*., in press). Further, the relationships between an undescribed extinct *Eudyptes* penguin from the NZ Chatham Islands (Tennyson & Millener, 1994; Millener, 1999), and the taxonomy of the extinct Hunter Island penguin (*Tasidyptes hunteri*) from Tasmania (van Tets & O'Connor, 1983) still remain unresolved or contested (Park & Fitzgerald, 2012).

Crucially, no previous analysis has considered all of these modern (including recently extinct) penguin taxa under a single phylogenetic framework. Similarly, most past phylogenetic approaches have relied primarily on morphology and/or short genetic sequences (e.g. Gavryushkina *et al*., 2017; *Figure 6*). We are now in the era of Next Generation Sequencing (NGS) and it is readily achievable to obtain whole mitochondrial genomes from extant and extinct taxa. Recent developments to bioinformatics platforms also enable us to better understand divergence events. For example, incorporating fossil-calibrated points, based on the geological and fossil record into a phylogenetic tree can provide direct estimates of when modern penguin clades may have diverged (Gavryushkina *et al*., 2017). As the phylogenetic interpretation of some of these potential calibration points has recently been re-considered (Degrange *et al*., 2018) or remain contentious, questions remain over the accuracy of recently dated penguin phylogenies (e.g. Gavryushkina *et al*., 2017; Frugone *et al*., 2018). As noted above, new discoveries, tools and data provide new opportunities to resolve and understand the timeframe of penguin cladogenesis.

Barriers to gene flow

Past population genetic studies have attempted to infer possible barriers to gene flow within single penguin taxa in the Southern Ocean. Several studies have sequenced thousands of single nucleotide polymorphisms across all *Pygoscelis* and *Aptenodytes* species (Clucas *et al*., 2016; Cristofari *et al*., 2016; Younger *et al*., 2017; Cristofari *et al*., 2018; Clucas *et al*., 2018). In

contrast, for *Eudyptes*, the only remaining widespread genus, just *E. chrysolophus chrysolophus* and *E. c. schlegeli* have been explored under a population genomic framework (Frugone *et al*., 2019), whereas the only other existing population genetic dataset for *Eudyptes* uses short sequences (Frugone *et al*., 2018), and is limited to only five of the eight recognised taxa. Most past studies focussed on ocean currents and the presence of the STF and APF as barriers to gene flow, arguing that these oceanographic features may have promoted lineage and species diversification in penguins, particularly within *Pygoscelis papua* (Clucas *et al*., 2018) and possibly for *Eudyptes moseleyi*/*E. filholi/E. chrysocome* (de Dinechin *et al*., 2009; Frugone *et al*., 2018) and *E. chrysolophus chrysolophus*/*E. c. schlegeli* (Frugone *et al*., 2018).

Population responses to global change

Several recent studies used multilocus genetic data to infer population genetic parameters and demographic histories within single penguin lineages (Clucas *et al*., 2014, Li *et al*., 2014; Pena *et al*., 2014; Younger *et al*., 2015a; Younger *et al*., 2015b; Clucas *et al*., 2016; Cristofari *et al*., 2016; Cristofari *et al*., 2018; Frugone *et al*., 2018; Dantas *et al*., 2019). By contrast, no study has yet attempted to present a multi-taxon synthesis of such data to test for the generality of such phenomena. A key goal in evolution and ecology should be to test for repeated responses of multiple taxa to ecosystem-wide change, to see whether such biological shifts are concerted or idiosyncratic (Wu, 2006; Smith *et al*., 2011; Marske *et al*., 2013). A multi-taxon demographic analysis is therefore crucial to help inform our understanding of key processes shaping the biodiversity and distribution of penguins in this rapidly warming part of the world (Thompson & Solomon, 2002).

Human-driven shifts in penguin biodiversity

New Zealand and its five sub-Antarctic islands (The Snares/Western Chain, Auckland Islands, Antipodes/Bounty Islands, Campbell Island, Chatham Islands) represent a notable hotspot for penguin diversity, with breeding colonies of seven extant species present across this region. Additionally, recent data suggest that prehistoric levels of penguin diversity may have been substantially higher in NZ, with several recent ancient DNA studies demonstrating human-driven shifts in the diversity of *Megadyptes* and *Eudyptula* penguins (see *Figure 4*). By contrast, comparatively little is known about the recent histories of the region's four *Eudyptes* species (*E. sclateri*, *E. robustus*, *E. pachyrhynchus*, *E. filholi*), three of which are endemic, or the possibility of additional recently extinct *Eudyptes* taxa (Tennyson & Millener, 1994; Millener, 1999). While *E. pachyrhynchus* is locally endemic to the oceanic-temperate forests on the west coast of the NZ South Island, the impacts of human arrival on demographic trends, range

contractions, or local species extinctions over the past several hundred years has never been explored under a molecular framework. It remains possible that the biogeographic ranges of sub-Antarctic *Eudyptes* penguins may have been more extensive than they are today (Worthy, 1997).

To date no study has expanded beyond the classic biogeographical and evolutionary examples for local island archipelagos, or indeed, explored how different Southern Ocean archipelagos may relate to each other over a circumpolar scale. Penguins provide an excellent model system to understand how the past histories of Southern Ocean islands may have shaped their recent evolution and biogeography. While penguins have the most extensive fossil record compared to any other Southern Ocean avian family, the resolution of the fossil record and genetic information independently from one another cannot be used to address all relationships within and between species. Clarifying these relationships in a comparative framework can provide new clues on how dynamic histories of Southern Ocean islands may be linked to penguin evolution and biogeography. The most appropriate way to do this is by using powerful mitochondrial and nuclear molecular approaches, in combination with fossil, geological and climate record data.

Genomic techniques as tools for studying penguin evolution

Every vertebrate contains genetic information for two genomes; a mitochondrial genome (mitogenome) and a nuclear genome. The haploid mitogenome is inherited through the maternal line, and is a powerful molecular marker of phylogenetic relationships among taxa (Moore & DeFilippis, 1997; Zink & Barrowclough, 2008), given its lack of recombination. While some studies are beginning to reconstruct species inter-relationships based on genome-wide nuclear data (e.g. Jarvis *et al*., 2014; Zhang *et al*., 2014a; Zhang *et al*., 2014b; Zhang *et al*., 2014c; Jarvis *et al*., 2015; Zhang, 2015; Kimball *et al*., 2019; Oliveros *et al*., 2019; Pan *et al*., in press), many studies rely on mitochondrial DNA (mtDNA) sequences to resolve phylogenies, especially when incorporating degraded ancient DNA (aDNA) data from recently extinct species.

Ancient DNA is the term used to describe genetic material preserved in degraded post-mortem biological remains. DNA extracted from ancient specimens is typically characterised by short fragment lengths and exogenous DNA from bacteria and other environmental sources (Knapp *et al*., 2012). Endogenous DNA is at such a low concentration that DNA extractions and pre-amplification steps must be undertaken in an ultra-clean environment to avoid contamination (Cooper & Poinar, 2000; Gilbert *et al*., 2005; Knapp *et al*., 2012; Skoglund *et al*., 2014). Many

studies are typically forced to use polymerase chain reaction (PCR) followed by Sanger sequencing to amplify aDNA in short fragments, usually about 100 base pairs (bp). Therefore, the majority of ancestral relationships that use aDNA have been studied using short fragments of mtDNA, such as the slow-evolving Cytochrome Oxidase subunit 1 (COI) region (for phylogenetic studies) or the high-mutating Control Region (CR) (for population-genetic studies). In the mid-2000s, aDNA research was boosted by the development of NGS technologies, which enabled sequencing of much larger aDNA libraries for a range of applications, including palaeogenomics, metagenomics and metabarcoding (Murray *et al*., 2013; Grealy *et al*., 2015; Heintzman *et al*., 2015; Slon *et al*., 2017; Seersholm *et al*., 2018; see also Cole & Wood, 2017b). Methods for efficiently retrieving very short DNA fragments (e.g. Dabney *et al*., 2013), and techniques such as hybridisation capture, which can enrich DNA libraries from a species of interest across an entire conserved mitogenome (e.g. Mitchell *et al*., 2014b) have been developed to maximise the potential of aDNA in the absence of a reference mitogenome. While some recent aDNA studies have successfully sequenced nuclear DNA from well-preserved specimens (Green *et al*., 2006; Noonan *et al*., 2006; Miller *et al*., 2008; Prüfer *et al*., 2014; Meyer *et al*., 2016; Dussex *et al*., 2019), conserved mtDNA (such as COI and mitogenomes) still remains a marker of choice for phylogenetic analyses of ancient specimens (e.g. Mitchell *et al*., 2014a; Mitchell *et al*., 2014b; Wood *et al*., 2014; Scofield *et al*., 2017; Boast *et al*., 2019; Knapp *et al*., 2019; Grealy *et al*., 2019). The rapid evolutionary rate of the CR makes it a good marker for recovering recent demographic history of populations and population structure in ancient specimens (e.g. Baker & Marshall, 1997; Dussex *et al*., 2015; Rawlence *et al*., 2015b; Rawlence *et al*., 2017a; Rawlence *et al*., 2017b).

Reduced representation sequencing of modern specimens can also be used to identify hundreds of thousands of single nucleotide polymorphisms (SNPs) in species without advanced genomic resources such as a reference genome (Baird *et al*., 2008). There are now a number of variations on the original method (reviewed in Andrews *et al*., 2016), but all rely on restriction enzymes to target conserved restruction enzyme cut sites at specific regions throughout the genome. This allows the same subset of the genome to be sequenced across tens to hundreds of individuals, and has now been adopted as the preferred method for conducting population genomic research in modern wild populations (Davey & Blaxter, 2010; Pimm *et al*., 2015; Allendorf, 2016), compared to previous microsatellite techniques (e.g. Moodley *et al*., 2015; Cole *et al*., 2016). Restriction sites can occur in both coding and non-coding regions of the genome, and will often be conserved across closely related taxa, which enables the sequence data to be used to answer a variety of evolutionary and demographic questions (Morin *et al*., 2004).

Aims and outline of this thesis

This thesis travels through time, spanning almost 20 million years of penguin evolution. Broadly, the aim of this thesis is to understand whether classic patterns of island evolution and biogeography extend to the islands in the circumpolar and vast Southern Ocean. While many penguin species are endemic to geologically young islands, other species are widespread, inhabiting islands with dynamic glacial or human colonisation histories. Given the large number of extant penguin species, combined with their extensive fossil record, this clade provides an excellent study system with which to assess how the history of islands have impacted biological diversification, biogeography and responses to environmental change.

This thesis is structured into four main research chapters. Each research chapter tests for possible links between the dynamic histories of Southern Ocean islands and the evolution of penguins. Taken together, these research chapters examine penguin evolution over temporal scales ranging from ancient to modern and spanning key geological, climatic and anthropogenic events and processes. The broad hypotheses for the four research chapters are as follows:

***Chapter 2*:** The emergence of geologically young islands in the Southern Ocean was crucial for the evolution of island-endemic penguin species, and highlights the importance of founder speciation.

***Chapter 3*:** Post-glacial warming of the Southern Ocean resulted in concerted, rapid southward expansions of numerous penguin species.

***Chapter 4*:** Human arrival to Pacific Southern Ocean islands drove the extinction of some endemic penguin taxa, and led to reduced genetic diversity in surviving lineages.

***Chapter 5*:** The extinct Southern Ocean endemic Hunter Island penguin is not a valid penguin taxon.

Chapter 2 clarifies the evolutionary history of modern penguins (16 mya to present), and identifies potential links between island emergence and the origins of island-endemic species. This study was performed using a combination of mtDNA sequences (COI and CR) and near-complete mitogenomes from ancient and modern specimens spanning all extant and recently extinct taxa. Phylogenetic analyses were undertaken using well-justified fossil calibration points, and results were compared to the geological record of island ages. In addition, this chapter describes two newly extinct penguin taxa (*Eudyptes* and *Megadyptes* genera) from the

NZ Chatham Islands, which were described based on aDNA sequence data and morphological characters. This chapter is published in *Molecular Biology and Evolution*.

Chapter 3 focusses on the post-glacial (1 mya to present) histories of all widespread penguin taxa among *Aptenodytes*, *Pygoscelis* and *Eudyptes* genera. All taxa have low genetic diversity including *Eudyptes chrysolophus chrysolophus* and *E. c. schlegeli* which probably represent incipient species. Concordant with post-glacial range expansions, all taxa that inhabit islands south of the LGM sea-ice extent experienced rapid post-glacial population expansions, while taxa that inhabit islands north of the LGM sea-ice extent have had stable/declining demographic histories. This was achieved by sequencing thousands of SNPs across seven *Eudyptes* taxa using blood samples of almost 300 individuals, and obtaining newly published SNP datasets of several hundred *Aptenodytes* and *Pygoscelis* penguins. This chapter is under revision in *Proceedings of the National Academy of Sciences*.

Chapter 4 builds on previous genetic studies focussing on how the arrival of Polynesians (780 years ago) and Europeans (250 years ago) in NZ may have shaped the biogeography of *Eudyptes* penguins. While only one *Eudyptes* penguin inhabits the NZ South Island today, six prehistoric taxa were detected, including a few individuals of the extinct Chatham Island *Eudyptes* penguin. Supplementing results from *Chapter 3*, demographic analyses suggest that *E. pachyrhynchus* has had a stable demographic history, possibly related to its habitat preference of isolated coastal forests. This was achieved using aDNA sequences (COI and CR) isolated from archaeological and subfossil bones, historical museum skins and modern blood samples from almost 200 specimens. This chapter is published in *Molecular Phylogenetics and Evolution.*

Chapter 5 clarifies the taxonomic identity of Tasmania's so-called 'extinct' *Tasidyptes hunteri*. This was achieved using aDNA sequences (COI) isolated from all known samples attributed to this genus. This study overturns the previous recognition of *T. hunteri* as a distinct and coherent species, instead showing that it consists of an artificial assemblage of three extant penguin species across two genera (*Eudyptes pachyrhynchus*, *E. robustus* and *Eudyptula novaehollandiae*). This chapter is published in the *Zoological Journal of the Linnean Society*.

Chapter 6 expands on the main themes of each research chapter, and synthesises the main results of this thesis. Future directions are discussed in terms of genomic techniques, outstanding questions about penguin evolution and biogeography and broad Southern Ocean island biology.

Taken together, this thesis provides the most comprehensive study on modern penguin evolution and genomics to date, and is the first multi-species study to link classic patterns of island biology and evolution to the diverse, circumpolar islands of the vast Southern Ocean.

Chapter 2

Mitogenomes uncover extinct penguin taxa and reveal island formation as a key driver of speciation

Publication details, contributions and acknowledgements

Parts of this chapter are published in:

Theresa L Cole[1,2], Daniel T Ksepka[3], Kieren J Mitchell[4], Alan JD Tennyson[5], Daniel B Thomas[6], Hailin Pan[7,8,9], Guojie Zhang[7,8,9], Nicolas J Rawlence[1], Jamie R Wood[2], Pere Bover[4,10], Juan L Bouzat[11], Alan Cooper[4], Steven R Fiddaman[12], Tom Hart[12], Gary Miller[13,14], Peter G Ryan[15], Lara D Shepherd[5], Janet M Wilmshurst[2,16] & Jonathan M Waters[1]. (2019). Mitogenomes uncover extinct penguin taxa and reveal island formation as a key driver of speciation. *Molecular Biology and Evolution.* 36(4): 784 – 797.

Author affiliations:

[1]Department of Zoology, University of Otago, PO Box 56, Dunedin 9054, Otago, New Zealand. [2]Manaaki Whenua Landcare Research, PO Box 69040, Lincoln 7640, Canterbury, New Zealand. [3]Bruce Museum, Greenwich, 06830, Connecticut, United States of America. [4]Australian Centre for Ancient DNA, School of Biological Sciences, University of Adelaide, Adelaide 5005, South Australia, Australia. [5]Museum of New Zealand Te Papa Tongarewa, PO Box 467, Wellington 6140, Wellington, New Zealand. [6]School of Natural and Computational Sciences, Massey University, Auckland 0632, Auckland, New Zealand. [7]State Key Laboratory of Genetic Resources and Evolution, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming 650223, China. [8]China National Genebank, BGI-Shenzhen, Shenzhen 518083, Guangdong, China. [9]Centre for Social Evolution, Department of Biology, Universitetsparken 15, University of Copenhagen, Copenhagen DK2100, Denmark. [10]ARAID Foundation, IUCA-Grupo Aragosaurus, Universidad de Zaragoza, Zaragoza 50009, Spain. [11]Department of Biological Sciences, Bowling Green State University, Bowling Green 43403, Ohio, United States of America. [12]Department of Zoology, University of Oxford, 11a Mansfield Road, Oxford OX1 3SZ, United Kingdom. [13]Division of Pathology and Laboratory Medicine, University of Western Australia, Crawley 6009, Western Australia, Australia. [14]Institute for Marine and Antarctic Studies, University of Tasmania, Hobart 7001, Tasmania, Australia. [15]DST-NRF Centre of Excellence, FitzPatrick Institute of African Ornithology, University of Cape Town, Rondebosch 7701, Republic of South Africa. [16]School of Environment, The University of Auckland, Private Bag 12019, Auckland 1142, Auckland, New Zealand.

Author contributions:

Theresa L Cole (95%) and Jonathan M Waters (5%) conceived and designed the study. Nicolas J Rawlence (50%), Theresa L Cole (40%) and Jamie R Wood (10%) sampled specimens from museums. Tom Hart (20%), Peter G Ryan (20%), Juan L Bouzat (20%), Gary Miller (20%) and Lara D Shepherd (20%) collected blood samples. Theresa L Cole (70%), Nicolas J Rawlence (25%) and Jamie R Wood (5%) undertook laboratory work for aDNA COI and CR sequencing. Theresa L Cole (75%), Kieren J Mitchell (10%), Pere Bover (8%), Jamie R Wood (5%) and Lara D Shepherd (2%) undertook laboratory work for aDNA mitogenome sequencing. Hailin Pan (90%), Theresa L Cole (5%) and Steven R Fiddaman (5%) undertook laboratory work for modern DNA Illumina sequencing. Theresa L Cole (85%), Kieren J

Mitchell (6%), Jamie R Wood (4%), Daniel T Ksepka (3%) and Pere Bover (2%) analysed the DNA data. Daniel T Ksepka (40%), Daniel B Thomas (40%), Theresa L Cole (15%) and Kieren J Mitchell (5%) compiled and analysed the fossil calibrations. Alan JD Tennyson (80%), Daniel T Ksepka (10%), Daniel B Thomas (5%) and Theresa L Cole (5%) analysed the morphological data and wrote the taxonomic descriptions. Theresa L Cole (65%), Guojie Zhang (10%), Jonathan M Waters (8%), Alan Cooper (5%), Janet Wilmshurst (5%), Jamie R Wood (5%) and Lara D Shepherd (2%) contributed to sequencing and laboratory costs. All co-authors helped to write the manuscript. Daniel T Ksepka, Kieren J Mitchell, Alan JD Tennyson and Daniel B Thomas contributed equally and are joint second author on the published article.

Acknowledgements:

We thank Matt Rayner for measurements, Vanesa de Pietri, Paul Scofield, Emma Burns and Trudi Webster for access to collections, Jean-Claude Stahl for photographs, Ewan Fordyce, James Scott, Nick Mortimer, Michael Knapp, Bill Lee, Paul Scofield and Trevor Worthy for discussions and Sam Petersen and Andy Black and Yves Cherel for field assistance and sample collection. We thank several anonymous reviewers for comments that improved the manuscript. Animal ethics came from the University of Otago, University of Oxford, University of Cape Town, Zoological Society of London and BGI. Permits were provided by the Government of South Georgia and the South Sandwich Islands and the South African National Antarctic Programme. This work was supported by the Marsden Fund (UOO1112, LCR1102), BGI, Australian Research Council, a National Science Foundation Award (DEB 1556615), a Rutherford Discovery Fellowship (RDFMNZ1201), Landcare Research and the University of Otago. Theresa L Cole was funded by an Otago University Postgraduate Scholarship.

Abstract

The emergence of islands has been linked to spectacular radiations of diverse organisms. Although penguins spend much of their lives at sea, they rely on land for nesting, and a high proportion of extant species are endemic to geologically-young islands. Islands may thus have been crucial to the evolutionary diversification of penguins. We test this hypothesis using a fossil-calibrated phylogeny of mitogenomes from all extant and recently extinct penguin taxa. Our analysis suggests that numerous island-endemic penguin taxa diverged following the formation of their islands, including *Spheniscus mediculus* (Galápagos Islands), *Eudyptes moseleyi* (Gough Island), *E. robustus* (The Snares) and *E. chrysolophus schlegeli* (Macquarie Island). Our analysis also reveals two new recently extinct island-endemic penguin taxa from NZ's Chatham Islands: the Chatham Island *Eudyptes* penguin (*Eudyptes warhami* sp. nov.) and a dwarf subspecies of the yellow-eyed penguin (*Megadyptes antipodes richdalei* ssp. nov.). *Eudyptes warhami* diverged from the Antipodes/Bounty Islands *E. sclateri* between 1.1 – 2.5 million years ago, shortly after the emergence of the Chatham Islands (approximately three million years ago). This new finding of recently-evolved taxa on this young archipelago provides further evidence that the radiation of penguins over the last five million years is linked to island emergence. Mitogenome analyses of all penguin species and the discovery of two newly extinct taxa highlight the importance of island formation in the diversification of

penguins, as well as the extent to which anthropogenic extinctions have affected island-endemic taxa across the Southern Hemisphere's isolated archipelagos.

Introduction

Biologists have long considered oceanic islands as natural laboratories for evolutionary studies (Darwin, 1859), with archipelago formation underpinning dramatic biological radiations in many remote regions of the globe (Shaw & Gillespie, 2016; see *Figure 1*). In particular, numerous studies have highlighted island isolation as a crucial prerequisite for species formation and adaptive radiation of island species (Darwin, 1859; Cowie & Holland, 2006; Losos & Ricklefs, 2009; Bacon *et al*., 2012). Soon after emergence, islands (whether volcanic or tectonic in origin) can be rapidly colonised by diverse arrays of dispersing taxa (Fleisher *et al*., 1998; Gathorne-Hardy & Jones, 2000; Mendelson & Shaw, 2005; Gillespie *et al*., 2012), presenting unique opportunities for local adaptation and diversification of species (Waters *et al*., 2013). The resultant island-endemic taxa can also be particularly prone to extinction (Shaw & Gillespie, 2006; Cowie & Holland, 2006; Wood *et al*., 2017).

Penguins are iconic flightless marine birds that inhabit all major landmasses and many islands in the Southern Hemisphere (*Figure 5*). Approximately 20 extant species are recognized, with some debate over species boundaries between recently-diverged populations (Frugone *et al*., 2018; Frugone *et al*., 2019; Mays *et al*., in press). The group has a rich fossil record extending back more than 60 million years (Slack *et al*., 2006), with over 50 extinct species documented (Ksepka *et al*., 2012, Simpson, 1971a; Myrcha *et al*., 1990; Clarke *et al*., 2003; Jadwiszczak, 2006; Walsh & Suárez, 2006; Ksepka & Ando, 2011; Park & Fitzgerald, 2012; Thomas & Ksepka, 2013; Mayr *et al*., 2017; Richards, 2019). Several phylogenetic studies have attempted to pinpoint the timing, and thereby the drivers, of penguin diversification (Baker *et al*., 2006; Ksepka *et al*., 2006; Subramanian *et al*., 2013; Gavryushkina *et al*., 2017). Previous studies identified circumpolar ocean currents and Antarctic cooling as key drivers of penguin evolution and biogeography (Baker *et al*., 2006; de Dinechin *et al*., 2009; Frugone *et al*., 2018). However, one third of all extant penguin species are endemic to geologically young islands (Maund *et al*., 1988; Gamble & Morris 1989; Adamson *et al*., 1996; Campbell *et al*., 2008; Sinton *et al*., 2018), suggesting that founder speciation may also have played an important role in recent penguin cladogenesis. An alternative explanation is that island-endemic penguins represent relictual populations of formerly more widespread species.

In this study, we test these competing hypotheses using 41 near-complete mitogenomes, representing all extant and recently extinct penguin taxa. By using well-justified fossil

calibrations, we resolve the timing and mechanisms of modern penguin diversification. We overlay our phylogeny using geological data of island ages, and demonstrate that several recent penguin divergence events correlate with the formation of islands, providing a new model for understanding penguin evolution. Furthermore, we describe two new recently extinct penguin taxa from the NZ Chatham Islands, demonstrating that while islands have been key in many recent penguin speciation events, the resulting restricted distributions have also made such lineages particularly susceptible to recent anthropogenic extinction.

Materials and Methods

DNA extraction, amplification and sequencing of historical samples

Historical skin samples from *Eudyptes filholi*, *E. robustus*, *E. sclateri*, *E. chrysolophus chrysolophus* and *E. c. schlegeli* were obtained from the National Museum of New Zealand Te Papa Tongarewa (NMNZ). 37 Holocene fossil and archaeological bones from the Chatham Islands which were identified as *Eudyptes* based on morphology, and two each of *Megadyptes antipodes* and *M. waitaha* were sourced from NMNZ, Canterbury Museum, and the Auckland War Memorial Museum (see *Supplementary Table 1*). To avoid duplicate sampling of individuals, either left or right elements were sampled from any one site, or bones were sampled from different stratigraphic units within a site. DNA extractions were performed following rigorous aDNA protocols as suggested by Cooper & Poinar (2000) at four purpose-built aDNA laboratories, depending on the sample type. These included (1) the Department of Zoology, University of Otago, Dunedin, following Rohland *et al*. (2010) for bone or Rawlence *et al*. (2015a) for historical museum skins; (2) Manaaki Whenua Landcare Research (MWLR), Lincoln, following Thomson *et al*. (2014) for bone; (3) NMNZ, Wellington, following the manufacturer's protocol from the Qiagen DNeasy® Tissue Extraction Kit (Qiagen Inc., Chatsworth, California, USA) for historical museum skins; and (4) the Australian Centre for Ancient DNA (ACAD), Adelaide, following Brotherton *et al*. (2013) for bone.

For species identification, we followed Boessenkool *et al*. (2008), Cole *et al*. (2018; see *Chapter 5*) and Cole *et al*. (2019a; see *Chapter 4*). We amplified approximately 499 bp of the COI gene (which was achieved through four overlapping 140 – 164 bp regions), 131 bp of the CR in *Eudyptes*, and approximately 402 bp of the CR in *Megadyptes* (which was achieved through two overlapping 229 – 255 bp regions). For primer details of both the COI and the *Eudyptes* CR refer to Cole *et al*. (2018; *Chapter 5*) and Cole *et al*. (2019a; *Chapter 4*), respectively. PCRs in a total volume of 12.5 μL each were performed using 2 mg/mL BSA (Sigma), 1 X PCR buffer, 2 mM $MgSO_4$, 80 μM dNTPs, 0.4 μM of each primer, 0.625 U HiFi

Platinum Taq (Invitrogen) and 1 μL DNA on a BIO-RAD MyCycler thermal cycler as follows: 94°C for 3 min; 55 cycles of 94°C for 30 s, 55°C for 30 s and 68°C for 45 s, followed by 68°C for 10 min. PCR products were purified using SPRIselect (Beckman Coulter, Inc., Indianapolis, IN, USA) and sequenced at the MWLR sequencing facility in Auckland on an Applied Biosystems 3500xL Genetic Analyzer. Contiguous sequences of COI and CR were assembled using Geneious v8.1.8 (Biomatters; Kearse *et al*., 2012) from high quality bidirectional reads, and checked manually. Due to aDNA damage, when inconsistency between sequences from a given individual was observed (e.g. G-A and C-T transitions), additional PCR and bidirectional sequencing were conducted and a majority rule consensus was applied (Brotherton *et al*., 2007).

Initial phylogenetic trees were created using BEAST v2.4.7 (Bouckaert *et al*., 2014) with a relaxed log-normal clock and Yule speciation model, 100 million markov chain monte carlo (MCMC) generations sampling tree parameters every 1000 generations, and a burnin of 10%. Analyses were run in triplicate and combined using Log Combiner v2.4.7. We used the Akaike Information Criterion (AIC) (Akaike, 1974) in JmodelTest2 (Darriba *et al*., 2012) to determine the most appropriate model of sequence evolution (Jukes Cantor for all genetic markers). We created two maximum-clade credibility phylogenies using the COI sequences, and one maximum-clade credibility phylogeny using the *Megadyptes* CR sequences (*Figures 7 - 8* and *Supplementary Figures 1-2*). As is typical with aDNA, our data contained some missing sequence data. We therefore constructed the first phylogeny which contained all COI sequences, and a second phylogeny that contained only samples with three to four of the four overlapping fragments. In addition, we included one sequence representative of all extant penguin species with a wandering albatross (*Diomedea exulans*) sequence as the outgroup. For the *Megadyptes* phylogeny we obtained 20 and 25 *M. antipodes* and *M. waitaha* sequences from GenBank, and one *Eudyptes chrysocome* sequence as the outgroup. Refer to *Supplementary Table 2* for sequences obtained from GenBank. We used PopArt (Leigh & Bryant, 2015) to create a minimum spanning haplotype network (Bandelt *et al*., 1999) that included all *Eudyptes* CR sequences.

Enriched mitochondrial genomes from sub-fossil bones and museum skins

While nuclear genomes have recently been used to reconstruct the phyogenomic history of penguins (Pan *et al*., in press), it remains difficult and prohibitively expensive to obtain full nuclear genomes from degraded museum skins or Holocene fossil and archaeological bones. Therefore, we obtained near-complete mitogenomes from all modern penguin taxa for our phylogenetic reconstruction.

One to two Illumina libraries were created for *Eudyptes filholi*, *E. robustus*, *E. pachyrhynchus*, *E. sclateri*, *E. chrysolophus schlegeli*, *E. warhami* (*cf. E*. clade X; see *Chapter 4*), *Megadyptes antipodes*, *M. waitaha* and *M. a. richdalei* (see results) following Meyer and Kircher (2012), but using truncated adapters with unique 7-mer barcode sequences and a partial uracil-DNA-glycosylase treatment (Rohland *et al*., 2015). We used real-time polymerase chain reaction (rtPCR) to determine the appropriate number of cycles to amplify each library (see Carøe *et al*., 2018). Two 10 µL reactions were run per library containing 1 µL of a 1:5 dilution of post-BSA product, 1 X HiFi PCR buffer, 2 mM $MgSO_4$, 0.25 mM dNTPs, 0.2 µM IS7 and IS8 primers (see Meyer & Kircher, 2012), 0.004 X ROX (Life Tech), 0.2 X SYBR (Life Tech), 0.56 M dimethyl sulfoxide (Sigma-Aldrich), and 0.2 U of Platinum Taq HiFi DNA polymerase (ThermoFisher). The rtPCRs were run on a LightCycler 96 (Roche) as follows: 94°C for 6 min; 40 cycles of 94°C for 30 s, 60°C for 30 s and 68°C for 40 s; and a high-resolution melt (95°C for 1 min, 40°C for 1 min then a ramp from 65 – 97°C). Each library was divided into 8 X 25 µL PCRs containing 3 µL of post-BSA library product, 1 X HiFi PCR buffer, 2 mM $MgSO_4$, 0.25 mM dNTPs, 0.2 µM IS7 and IS8 primers and 0.5 U Platinum Taq DNA HiFi polymerase (ThermoFisher). PCRs were run in a heated-lid thermal cycler as follows: 94°C for 6 min; 10 – 22 cycles (determined by rtPCR) of 94°C for 30 s, 60°C for 30 s and 68°C for 40 s; followed by 68°C for 10 min.

Libraries were enriched for avian mtDNA (except for the CR) with commercially synthesised biotinylated 80-mer ribosomal ribonucleic acid (RNA) baits (Arbor Biosciences, MI, USA), designed from published mitogenome sequences for 27 modern birds, (including *Eudyptes* and *Megadyptes*; see Mitchell *et al*., 2014b). DNA-hybridisation enrichment was performed according to the manufacturer's recommendations (myBaits protocol v1), except the incubation step, which we extended to 44 h (2 h at 60°C, 12 h at 55°C, 12 h at 50°C and 17 h at 55°C). After washing the bound DNA, the baits and DNA library were eluted in PCR master mix, which was then divided into 5 X 25 µL reactions comprising 1 X AmpliTaq Gold Buffer, 2.5 mM $MgSO_4$, 0.25 mM dNTPs, 0.4 µM indexed full-length adapter primers (IS4 and indexing primer; see Meyer & Kircher, 2010) and 1.25 U AmpliTaq Gold DNA Polymerase (Thermo Fisher). PCRs were run in a heated-lid thermal cycler as follows: 94°C for 6 min; 15 cycles of 94°C for 30 s, 60°C for 30 s and 72°C for 45 s; 72°C for 10 min. Each amplified library was diluted to 2 nM and run on an Illumina MiSeq using 2 X 150 bp paired-end sequencing chemistry.

Reads were demultiplexed using 'sabre' (http://github.com/najoshi/sabre) (default parameters: no mismatches allowed). Adapter sequences were removed and paired-end reads were merged

using AdapterRemoval v2.1.2 (Schubert *et al*., 2016). Low quality bases were trimmed (<Phred20 --minquality 4) and merged reads containing <25 bp were discarded (--minlength 25). Read quality was visualised using fastQC v0.10.1 (http://www.bioinformatics.bbsrc.ac.uk/projects/fastqc) before and after trimming to make sure the trimming was efficient. Collapsed reads from *Eudyptes sclateri* (AD302), *E. warhami*/*E*. clade X (AD342) and *Megadyptes antipodes richdalei* (ACAD12997) were mapped against the *Eudyptes chrysocome* mitogenome (GenBank accession AP009189) using BWA v0.7.8 (Li & Durbin, 2009) (aln -1 1024, -n 0.01, -o 2). Reads with a mapping quality Phred score >30 were selected using the SAMtools v.1.4 (Li *et al*., 2009) view command (-q 30) and duplicate reads were discarded using 'FilterUniqueSAMCons.py' (Kircher, 2012). We created 50% (majority-rule) consensus sequences in Geneious. Collapsed reads from all other samples were mapped to these three new reference sequences, as described above. Preliminary taxon assignments were made based on the reference to which the most reads mapped, and 50% consensus sequences were created based on the closest matching reference (*E. sclateri* for all extant *Eudyptes* samples, *E. warhami*/*E*. clade X for all *E. warhami*/*E*. clade X samples and *Megadyptes antipodes richdalei* for all *Megadyptes* samples). All reads for each sample were then re-mapped to their respective consensus, as described above, and final higher-stringency consensus sequences were created: 85% majority for all samples, except *Eudyptes warhami*/*E*. clade X which were called at 100% majority. A nucleotide was only called for sites with >3X coverage depth, and sites with insufficient read depth were called as 'N'. For two *Megadyptes antipodes richdalei* samples (AD161 and AD88) we were able to combine two sets of reads that were created from two Illumina libraries to create one sequence per sample. For these samples, there were no ambiguities between reads that overlapped. We used MapDamage v.2.0.6 (Jónsson *et al*., 2013) to ensure that damage patterns in our data were consistent with authentic aDNA (see *Supplementary Figures 3 – 6* and *Supplementary Table 3*).

Mitogenomes from contemporary blood

We created mitogenomes from blood samples of *Aptenodytes patagonicus* (Fortuna Bay, South Georgia)*, Eudyptes moseleyi* (Amsterdam Island)*, E. chrysolophus chrysolophus* (Marion Island) and *E. c. schlegeli* (Green Gorge, Macquarie Island) at BGI, Hong Kong. DNA was extracted at either BGI, Hong Kong or the University of Oxford, Oxford using a HiPure Blood DNA Midi Kit II or the Qiagen DNeasy® Tissue Extraction Kit, respectively. We constructed 250 bp insert size libraries and performed whole-genome paired-end sequencing on a BGISeq 500 platform with 150 bp read lengths (see Pan *et al*., in press). Mitogenomes were *de novo* assembled using SOAPdenovo-Trans (Xie *et al*., 2014) with approximately 6 Gigabases of data

for each species. We raised the linkage support during the scaffolding step to improve accuracy and to avoid connections between mtDNA reads and nuclear mtDNA segments (NUMTs). We used a hidden Markov model method (Krogh *et al*., 1994; Durbin *et al*., 1998; Wheeler & Eddy, 2013) to filter candidate mtDNA sequences and remove potential false positive mitochondrial scaffolds, NUMTs or sequences that did not map to avian mitochondrial genes, which we considered as contamination.

Phylogenetic analysis of mitogenomes

We aligned all mitogenomes, including 15 additional mitogenomes obtained from GenBank (see *Supplementary Table 2*) using the MUSCLE algorithm implemented in Geneious R8. As most of the mitogenomes were not enriched for the CR, we excluded this region from genomes sequenced on the BGISeq 500 platform or those that had been downloaded from GenBank. We provided PartitionFinder v2.1.1 (Lanfear *et al*., 2016) with an input of 30 regions: for each protein coding gene, a region corresponding to every third position counting from the first position (e.g. position 1, 4, 7, etc.) and a region corresponding to every third position counting from the second position (e.g. position 2, 5, 8, etc.); 12S ribosomal RNA; 16S ribosomal RNA; concatenated transfer RNAs; and 108 bp of concatenated non-coding regions (*Supplementary Table 4*). The two regions we defined for each protein coding gene correspond to the first and second codon positions, except for the last 178 bp of the ND3 gene, which are frameshifted by one position due to a single nucleotide insertion. Regions were defined by aligning our sequences with the *E. chrysocome* mitogenome (GenBank accession AP009189) using Geneious, and extracting and concatenating the gene regions for downstream analyses. Optimal partitioning schemes were chosen based on the Bayesian information criterion (BIC) without a maximum-likelihood (ML) starting tree (Lanfear *et al*., 2012; Guindon *et al*., 2010) (*Supplementary Table 4*). We used Geneious to calculate genetic divergences among taxa using both the entire alignment (*Supplementary Table 5*) and excluding any positions with missing data in the alignment (*Supplementary Table 6*). We analysed all partitioned alignments in BEAST to infer the topology of penguins under a phylogenetic framework (*Supplementary Figure 7* and *8*). We explored two rooting strategies, using Steller's albatross (*Phoebastria albatrus*) as the outgroup (a member of Procellariiformes, commonly recognised as the sister order to penguins; Ksepka *et al*., 2006) and rooting the phylogenetic tree between *Aptenodytes*/*Pygoscelis* and all other penguins (*Supplementary Figure 7* and *8*). All initial phylogenetic analyses were implemented with a Yule process speciation prior, under a single lognormal relaxed clock model. All phylogenetic analyses were conducted via the Cyber

Infrastructure for Phylogenetic RESearch (CIPRES) Science Gateway v3.3 (Miller *et al*., 2010).

Phylogenetic analysis using fossil calibrations

To calibrate mitogenomic evolution, we constrained the ages of five key nodes based on fossil penguins. Details and justifications for each node, and the parameters implemented in BEAST are described below.

Fossil calibrations

Calibrated Node 1

Spheniscifformes-Procellariiformes

Taxon

Waimanu manneringi

Specimen

CM zfa35, partial skeleton (holotype)

Phylogenetic justification

Phylogenetic analyses universally support *Waimanu* being along the stem penguin lineage (e.g. Slack *et al*., 2006; Ksepka *et al*., 2006; Gavryushkina *et al*., 2017; Chávez Hoffmeister, 2014). CM zfa35 is the only published specimen of *W. manneringi* and all previous phylogenetic analyses have been based on this material.

Minimum age constraint

60.5 Ma

Maximum age constraint

72.1 Ma

Prior distribution

Lognormal

Mean

3.7 (in real space)

Standard deviation

1.0

Offset

60.5

Age justification

Biostratigraphic evidence, specifically the ranges of *Hornibrookina teuriensis* and *Chaismolithus bidens* indicate the minimum possible age of the type locality is 60.5 Ma (Cooper, 2004; Slack *et al*., 2006; Ogg *et al*., 2008). The maximum age constraint is based on the lower bound of the Maastrichtian Stage. Maastrichtian marine vertebrate sites have yielded several diving birds, including the Northern Hemisphere hesperornithids and the Southern Hemisphere *Neogaeornis*, *Vegavis* and *Polarornis*, indicating preservation potential for marine diving birds globally, which is specifically within the geographic range of modern penguins.

Calibrated Node 2

Crown Spheniscidae

Taxon

Madrynornis mirandus

Specimen

MEF-PV 100, partial articulated skeleton (holotype)

Phylogenetic justification

M. mirandus was originally considered to be closely related to *Eudyptes* (Acosta Hospitaleche *et al*., 2007; Ksepka & Clarke, 2010), and later considered to possibly represent the sister taxon to crown Spheniscidae (Chávez Hoffmeister, 2014; Chávez Hoffmeister *et al*., 2014). Re-examination of the holotype has revealed new character evidence and the most recent phylogenetic analysis suggested that *M. mirandus* is instead more closely related to *Spheniscus* and *Eudyptula*, though support was weak for this hypothesis (trees placing the fossil with *Eudyptes* were only one step longer; Degrange *et al*., 2018). Nevertheless, seven synapomorphies support crown status for *M. mirandus*, most compellingly the widely separated fossa temporalis, elongate processus retroarticularis, and small foramen ilioischiadicum (Degrange *et al*., 2018). Given the strong evidence that *M. mirandus* is a crown penguin and

the uncertainty over the precise placement of this taxon, we use *M. mirandus* as a calibration for the penguin crown.

Minimum age constraint

9.7 Ma

Maximum age constraint

25.2 Ma

Prior distribution

Uniform

Age justification

The single known specimen of *M. mirandus*, comprising most of a skeleton, was collected from the Entrerriense sequence of the Puerto Madryn Formation (Acosta Hospitaleche *et al*., 2007). This sequence was deposited at 10.0 ± 0.3 Ma based on strontium isotope dating obtained from fossil molluscs (Scasso *et al*., 2001). The maximum age is based on the age of the upper boundary of Kokoamu Greensand of NZ, units that together have yielded a large number of penguin specimens spanning a wide range of body sizes and representing at least five different species, all of which are demonstrably stem taxa. Because the boundary between Kokoamu Greensand and the overlying Otekaike Limestone likely occurs near the upper Whaingaroan/Duntroonian boundary, we arbitrarily use the Whaingaroan/Duntroonian boundary (25.2 Ma) here in the absence of a more refined date. This maximum age for the crown is consistent with the observations that Oligocene units in Australia and South America have also yielded exclusively stem penguins (no fossil Oligocene penguins have yet been reported from Antarctica or Africa).

Calibrated Node 3

Spheniscus-Eudyptula

Taxon

Spheniscus muizoni

Specimen

MNHN PPI 147, partial skeleton (holotype)

Phylogenetic justification

Göhlich (2007) listed a number of diagnostic characters supporting placement of *S. muizoni* in the genus *Spheniscus*. The placement of *S. muizoni* has been supported by several subsequent phylogenetic analyses (e.g. Ksepka & Clarke, 2010; Chávez Hoffmeister *et al.*, (2014). Unambiguous synapomorphies of *Spheniscus* preserved in the calibrating specimen include the straight proximal border of the fossa tricipitalis in ventral view (observed only in *Spheniscus* and some specimens of *Palaeospheniscus*) and extremely deep sulcus longitudinalis dorsalis medialis (observed only in *Spheniscus* and the otherwise very dissimilar *Aptenodytes*).

Minimum age constraint

9.2 Ma

Maximum age constraint

23.03 Ma

Prior distribution

Lognormal

Mean

4.4 (in real space)

Standard deviation

1.0

Offset

9.2

Age justification

The calibrating specimen was collected from the Cerro la Bruja locality of the Pisco Formation of Peru. In the description, an age estimate of 11 – 13 Ma for *S. muizoni* was based on general faunal divisions (Göhlich, 2007). However, subsequent stratigraphic work (Brand *et al*., 2011) provides a revised age of 9.2 Ma for Cerro la Bruja. The maximum age is set to the base of the Miocene (23.03 mya; Gradstein, 2012). This maximum encompasses (1) the well-studied outcrops of the South American early Miocene Gaiman Formation, known as 'Patagonia Formation' in older references, which have yielded abundant stem penguin fossils (e.g. *Palaeosphneiscus*, *Paraptenodytes*, *Eretiscus*) but no crown penguins; (2) Miocene localities

in Australia, all of which are interpreted as either stem penguins or too incomplete to be assigned to either the stem or crown (Park *et al*., 2016); (3) Miocene-Pliocene record of NZ, which has yielded crown penguins, none of which fall within the *Spheniscus*-*Eudyptula* clade; and (4) the African penguin record, which is limited to middle/late Miocene material of indeterminate status (Thomas & Ksepka, 2013) and several early Pliocene crown species, none of which fall within the *Spheniscus*-*Eudyptula* clade (Ksepka & Thomas, 2012).

Calibrated Node 4

Eudyptes-*Megadyptes*

Taxon

Eudyptes sp.

Specimen

NMNZ S.046318, partial skeleton

Phylogenetic justification

The strongly arched jugal bar in this specimen is a derived feature of *Eudyptes* (Daniel Thomas, personal communication). Although this feature also occurs in some species of *Pygoscelis*, tarsometatarsi also referred to this specimen have the derived condition of the foramen vasculare proximale medialis perforating the crista medialis hypotarsi (rather than exiting distal to the crest) supporting a position close to *Eudyptes* and ruling out a relationship with *Pygoscelis.*

Minimum age constraint

3.06 Ma

Maximum age constraint

25.2 Ma

Prior distribution

Lognormal

Mean

7.04 (in real space)

Standard deviation

1.0

Offset

3.06

Age justification

The calibrating *Eudyptes* sp. fossil is from the Late Pliocene Tangahoe Formation in the southern Taranaki region of the North Island of NZ (Naish *et al*., 2005). The Tangahoe Formation has been tightly constrained between 3.36 – 3.06 Ma and is therefore within the local Waipipian stage (3.7 – 3.0 Ma) and the international Piacenzian stage (3.6 – 2.58 Ma) (Naish *et al*., 2005; Raine *et al*., 2015). The formation was dated using magnetostratigraphic correlation with the Ocean Drilling Program Site 846 chronology, as well as the presence of Waipipian stage macrofossils and microfossils (Naish *et al*., 2005). *E. calauina* from the Horcón Formation of Chile is of similar age but is known from less complete material. Chávez Hoffmeister *et al*. (2014) recovered *E. calauina* within a polytomy that included all extant *Eudyptes* species. The Horcón Formation is considered Late Pliocene but no finer numerical dates are available for the horizon from which *E. calauina* is known. Thus, *E. calauina* may be slightly older or slightly younger than the Taranaki *Eudyptes* sp. Because many species of the *Eudyptes-Megadyptes* clade occur predominantly on islands with no pre-Holocene fossil records today, we used a conservative Oligocene maximum that follows the same justification as that for crown penguins as a whole.

Calibrated Node 5

Aptenodytes-Pygoscelis

Taxon

Pygoscelis calderensis

Specimen

SGO-PV 790, partial skull

Phylogenetic justification

P. calderensis was described from three partial skulls. The holotype skull preserves a very shallow temporal fossa, a derived feature which is present only in *Aptenodytes* and *Pygoscelis* penguins. The holotype skull also preserves a shelf of bone bordering the supraorbital salt gland

fossa, a second derived feature which occurs in *Pygoscelis* (as well as *Megadyptes* and *Eudyptes*) but is absent in *Aptenodytes*. Together, these two features support placement of *Pygoscelis calderensis* to at least to the stem of *Pygoscelis*.

Minimum age constraint

6.3 Ma

Maximum age constraint

25.2 Ma

Prior distribution

Lognormal

Mean

6.0 (in real space)

Standard deviation

1.0

Offset

6.3

Age justification

The calibrating specimen was collected from the Bahía Inglesa Formation. This formation contains several fossil-bearing horizons. The uppermost horizon, the Lechero Member, contains an ash layer which provides a potassium-argon age of 7.6 ± 1.3 Ma (Marquardt *et al*., 2000; Godoy *et al*., 2003). *P. calderensis* was collected from a phosphatic horizon several meters lower in the stratigraphic section. The dated ash layer, therefore, provides a hard minimum age for the fossil. Because both *Pygoscelis* and *Aptenodytes* occur predominantly in Antarctica and sub-Antarctic islands today, and the fossil record from Antarctica is very poor aside from the Eocene deposits of Seymour Island (Simpson, 1971b; Jadwiszczak, 2006; Myrcha, 2006), we use a conservative Oligocene maximum that follows the same justification as that for crown penguins as a whole.

Phylogenomic analysis

To calibrate our mitogenome phylogeny, we implemented a single lognormal relaxed clock model and constrained the ages of five key nodes (four when *Phoebastria albatrus* was excluded) based on the fossil penguins (as described above). To test the contribution of including the calibration stem penguin *Waimanu manneringi* we ran multiple phylogenetic tests using the alignment that sampled one individual per species, with *Phoebastria albatrus* as an outgroup. The tests used the following calibration points: (1) only *Waimanu manneringi* (see *Supplementary Figure 9C*); (2) all calibrations including *W. manneringi* (see *Supplementary Figure 9B*); and (3) all calibrations excluding *W. manneringi* (see *Supplementary Figure 9A*). The performance of each test was assessed by running the analysis without the data (e.g. priors only, see Warnock *et al*., 2015) to determine the relative contribution of the data and the priors to the posterior node age estimates. When *W. manneringi* was included, the effective priors for the crown nodes were much older than the initial specifications, yet when *W. manneringi* was removed the split between Sphenisciformes and Procellariiformes became unrealistically young, based on the fossil evidence (see *Supplementary Figure 9*). This conflict could be due to a rate slowdown in the large-bodied penguin clade (Ksepka & Phillips, 2015) and may be exacerbated by the long branch separating crown penguins from Procellariiformes (see *Supplementary Figure 9*). Therefore, we removed the outgroup *Phoebastria albatrus*, and ran the analysis on just Sphenisciformes without the *Waimanu manneringi* calibration point. Each test was run at a minimum of 50 million MCMC, and for each new alignment we re-ran Partition Finder and a single uncalibrated phylogenetic analysis was performed as above to determine the topology (especially for *Aptenodytes*/*Pygoscelis* placements, which remained as sister taxa in all analyses). To assess whether subspecies versus species status impacted the relationships among closely-related Sphenisciformes taxa (e.g. *Eudyptes chrysolophus chrysolophus*/*E. c. schlegeli*) we ran all analyses with a conservative approach, where we considered *E. c. chrysolophus*/*E. c. schlegeli* and two *Megadyptes* subspecies (see results) as conspecific (*Supplementary Figure 10*). To assess the impact of tree prior choice we ran all analyses (with and without *Phoebastria albatrus*) using either a birth-death or a calibrated Yule speciation process (*Supplementary Figure 10* and *11*). Results did not differ substantially, so we consider only the analyses that used the birth-death speciation prior (see *Figure 9C* and *Supplementary figure 10* and *11A*). After comparing the performance of each test, the final analyses were run using only Sphenisciformes, with four internal calibration points, including a uniform prior for crown penguins (*Figure 9C*). This final analysis was run for 100 million generations, and we sampled trees and parameter values every 1000 generations. Parameter values were monitored

and compared between chains in Tracer v1.6 (Rambaut *et al*., 2013) to ensure convergence and effective sample sizes (ESS) >200. We combined sampled trees and parameter values from each chain using Log Combiner. The first 10% of each chain was discarded as burn-in using TreeAnnotator v1.8.3. We visualised the maximum-clade credibility tree using FigTree v1.4.2. Sequences and calibrated tree files are available on GenBank (MK290241–MK290284) and/or Figshare (DOI: 10.6084/m9.figshare.c.4329029).

Systematic Palaeontology

We measured 12 elements from eight *Eudyptes* taxa (*n* = 87; *Supplementary Table 7*), and up to 23 elements from each *Megadyptes* taxa (*n* = 57; see below and *Supplementary Tables 8 – 12*) to examine possible morphological distinctions within and between *Eudyptes* and *Megadyptes* taxa (see below). Radiocarbon dates were obtained from terrestrial birds from the same localities as described material (Millener, 1999), and were re-calibrated using the Southern Hemisphere terrestrial calibration curve (Hogg *et al*., 2013) via OxCal v4.3.2 (Bronk Ramsey, 2017).

Results

Genetic evidence for two extinct penguin lineages from the Chatham Islands

We analysed 65 sub-fossil bones (*Supplementary Table 1*) from the Chatham Islands to test for the existence of an extinct *Eudyptes* species, as proposed by Tennyson and Millener (1994) and Millener (1999) based on morphological evidence. Most bones were poorly preserved but we obtained partial COI sequences from 22, and partial CR sequences from eight. Phylogenetic analyses of these data identified *E. sclateri* and two other distinct genetic lineages, one corresponding to *Eudyptes* clade X (see Cole *et al*., 2019a; *Chapter 4*), *Figure 7* and *Supplementary Figures 1 – 2*) and another within the *Megadyptes* genus (see *Figure 8* and *Supplementary Figure 12*). The latter discovery was unexpected, as the bones had appeared too small to belong to *Megadyptes* and thus had originally been identified as *Eudyptes*.

Figure 7. Bayesian phylogenetic tree of penguins, with high-quality aDNA COI sequences from Chatham Islands bone samples. Prehistoric samples have AD numbers. One bone from the Chatham Islands is identified as *Megadyptes* (purple), two are *Eudyptes sclateri* (orange) and 10 fall with *Eudyptes* clade X (red) from mainland NZ (discussed in Cole *et al*. 2019a; *Chapter 4*]). Posterior probabilities for the main clades are shown.

Mitogenomic phylogenetic analysis

Phylogenetic analysis of near-complete mitogenomes recovered a clade grouping the two largest-bodied and most polar-adapted penguin genera, *Aptenodytes* and *Pygoscelis*, as sister to all other extant penguins (see *Figure 9C* and *Supplementary Figures 8 – 12*). This basal split was also found in two recent studies (Subramanian *et al*., 2013; Gavryushkina *et al*., 2017), whereas others have recovered *Aptenodytes* as an independent lineage, sister to all other extant penguins (Bertelli & Gianni, 2005; Baker *et al*., 2006; Ksepka *et al*., 2006, Pan *et al*., in press). Our phylogeny agrees with previous studies in recovering *Spheniscus* + *Eudyptula* and *Eudyptes* + *Megadyptes* clades (Bertelli & Gianni, 2005; Baker *et al*., 2006; Ksepka *et al*., 2006). The Chatham Islands *Eudyptes* taxon is distinct, sister to *E. sclateri* (posterior probability: 1.0). The individuals recognised as *Eudyptes* clade X from mainland NZ (see Cole *et al*., 2019a; *Chapter 4*) belong to this newly-recognised species. Samples from the previously unrecognised 'dwarf' Chatham Islands *Megadyptes* taxon (marked in purple in *Figure 8*) form a monophyletic clade (posterior probability: 1 in *Figure 8*; posterior probability: 0.88 in *Figure 9*), though their precise phylogenetic relationship to other *Megadyptes* lineages could not be confidently resolved (*Figure 8*).

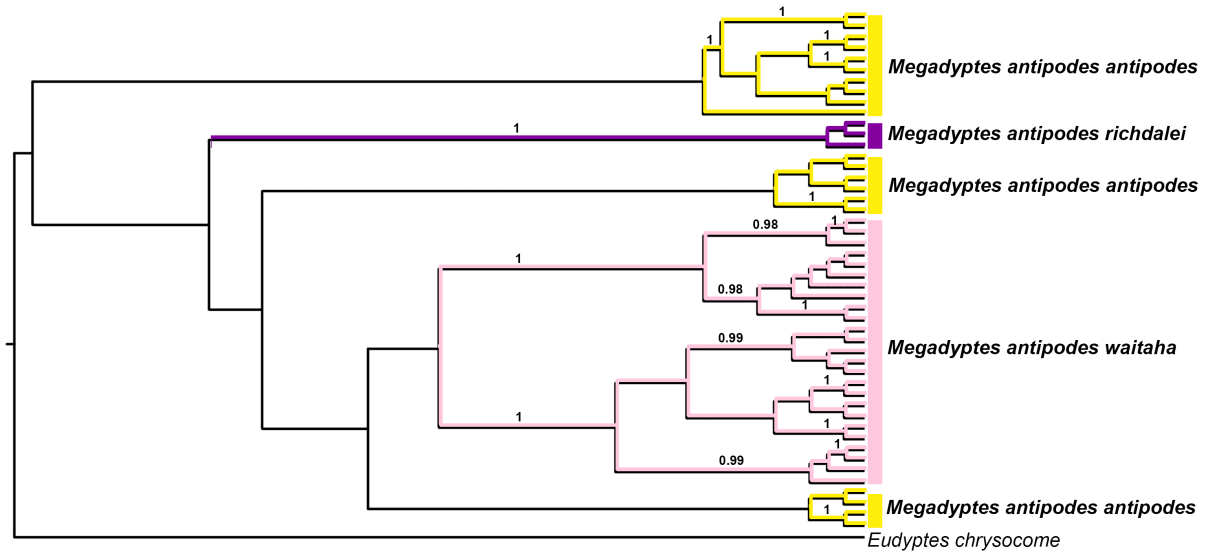


Figure 8. Bayesian phylogenetic tree of *Megadyptes* penguins with aDNA CR sequences from Chatham Islands bone samples. Posterior probabilities >0.95 are shown.

Pairwise distances and taxonomic status

To explore genetic divergences and potential taxonomic status of penguin lineages, we compared the mitogenomes in a pairwise matrix (see *Figure 9B* and *Supplementary Tables 5 – 6*). The Chatham Islands *Eudyptes* penguin and *E. sclateri* are 1.9% divergent, with 268 differential SNPs between the two species. In contrast we observed just 0.2% divergence between the closely-related (Christidis & Boles, 2008; Frugone *et al*., 2018; Frugone *et al*., 2019) *E. chrysolophus schlegeli* and *E. c. chrysolophus* lineages. Despite clear phenotypic differences (Warham, 1975; Shaughnessy, 1975; Woehler, 1995) it appears that these genetically similar lineages may still be in the earliest stages of diversification, supporting recent conclusions of Frugone *et al*. (2018) and Frugone *et al*., (2019); see also *Chapter 3*. Our data reveal more substantial divergences among three rockhopper penguin species (de Dinechin *et al*., 2009): *E. chrysocome* and *E. filholi* are 0.7% divergent, and both show 1.8% divergence relative to *E. moseleyi*, supporting recent conclusions of Frugone *et al*. (2018); see also *Chapter 3*. The recently proposed recognition of two *Eudyptula* taxa (Grosser *et al*., 2017) is supported by 2.9% divergence detected between *Eudyptula minor* and *E. novaehollandiae*.

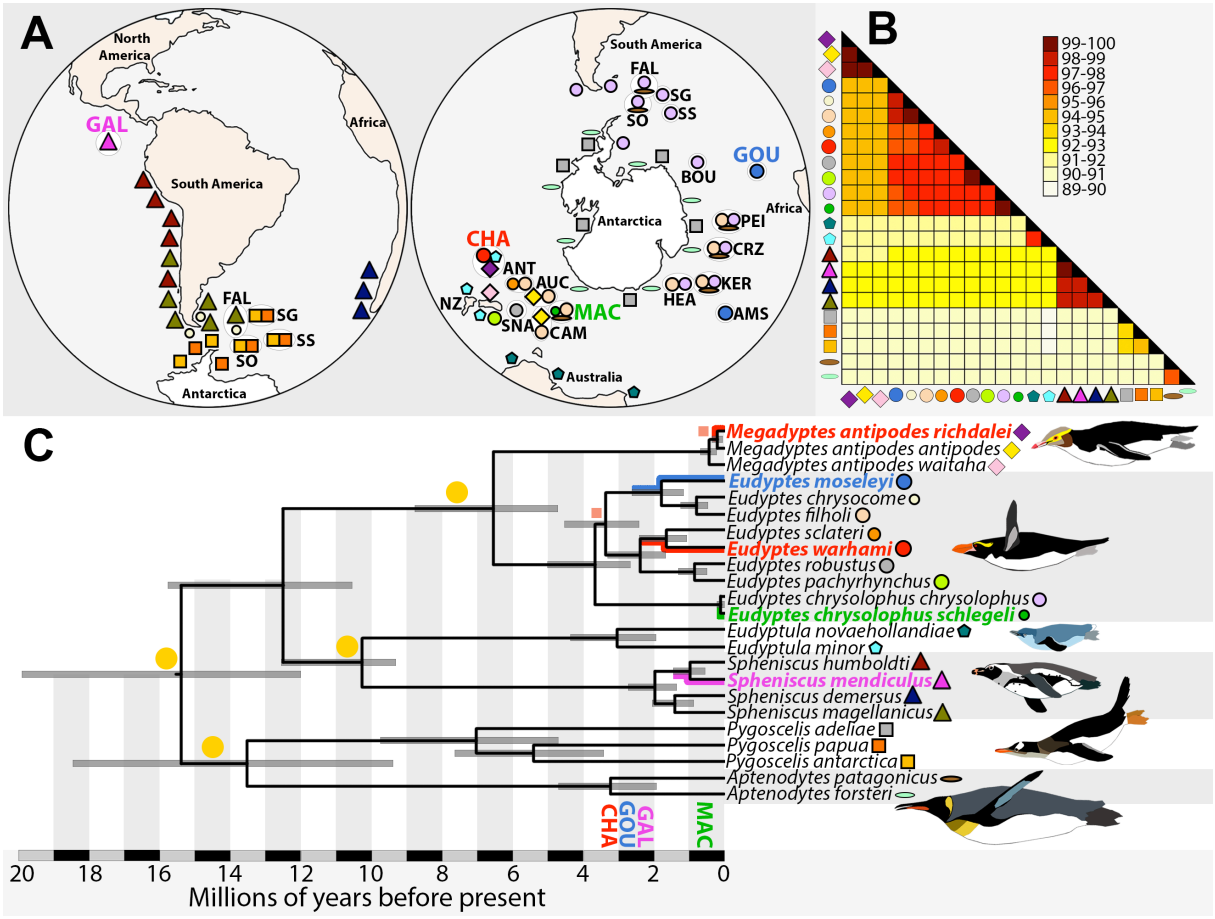


Figure 9. Breeding distributions and genomic relatedness between modern penguin taxa based on mitogenomes. **(A)** Maps of penguin breeding ranges, adapted from Ramos *et al*. (2018). Only the pre-human breeding range of *Megadyptes antipodes antipodes* is shown. GAL: Galápagos Islands; FAL: Falkland Islands; SG: South Georgia; SS: South Sandwich Islands; SO: South Orkney Islands; BOU: Bouvet; GOU: Gough Island; TDC: Tristan da Cunha; PEI: Prince Edward Islands; CRZ: Crozet Islands; KER: Kerguelen Islands; HEA: Heard Island; AMS: Amsterdam and St Paul Islands; MAC: Macquarie Island; CAM: Campbell Island; SNA: Snares; AUC: Auckland Islands; ANT: Antipodes Islands; CHA: Chatham Islands; NZ: New Zealand. **(B)** heatmap showing the percentage of pairwise genetic similarity calculated from 14,117 bp of 23 crown penguin mitogenomes (excluding all missing data); **(C)** dated phylogeny of penguins inferred from mitogenomes. Fossil calibrations are marked with large yellow circles. Posterior probabilities for all clades were >0.99, except those marked with orange squares (0.88 for *Megadyptes* and 0.74 for *Eudyptes*). 95% highest posterior denisty error bars are shown for each node. The divergence dates and the geological emergence of the Chatham Islands, Gough Island, Galápagos Islands and Macquarie Island are shown.

Megadyptes antipodes and the new Chatham Islands *Megadyptes* taxon had only 0.1% sequence divergence, and both were 0.3% divergent from *M. waitaha*. These values are substantially smaller than the divergences observed between some other sister species pairs of extant penguins (mean 2.2%, range 0.8 – 5.2%; *Supplementary Table 6*). In contrast, the Chatham Islands *Eudyptes* taxon is 1.9% divergent from *E. sclateri*, a value exceeding those between sister-pairs of several more widely accepted penguin species (e.g. *E. robustus* and *E. pachyrhynchus* are 0.8% divergent; see also *Chapter 3*).

Calibrated phylogenetic analysis

Our divergence estimates of major phylogenetic clades are consistent with the proposed Miocene origin of crown penguins (Subramanian *et al*., 2013; Gavryushkina *et al*., 2017) and show that a large proportion of penguin species (16 out of 23 studied taxa) have diverged over the past two million years (see *Figure 9C* and *Supplementary Table 13*). While some divergence-time estimates substantially predate or postdate island formation, the majority of island-endemic penguin taxa (The Chatham Islands *Eudyptes* on the Chatham Islands, *E. chrysolophus schlegeli* on Macquarie Island, *E. moseleyi* on Gough Island, the Chatham Islands *Megadyptes* taxon on the Chatham Islands, *Spheniscus mendiculus* on the Galápagos Islands and possibly *Eudyptes robustus* on The Snares) diverged from their respective sister taxa following the emergence of the islands they are endemic to (see *Figure 9C* and *Supplementary Figures 9 – 11* for alternative calibration approaches).

Systematic palaeontology

Morphological observations (*Supplementary Tables 7 – 12*) combined with our molecular results (*Figures 7, 9* and *10*) support recognition of the Chatham Islands *Eudyptes* penguin (originally proposed by Tennyson and Millener, 1994) as a distinct species (*Figure 10*) and the dwarf Chatham Islands *Megadyptes* penguin as a subspecies of the extant yellow-eyed penguin (*Figure 11*). Comparison of bone sizes and proportions are presented under the caveat that currently-available sample sizes are too low for robust conclusions using parametric tests.

Aves Linnaeus, 1758

Sphenisciformes Sharpe, 1891

Eudyptes Vieillot, 1816

Eudyptes warhami, nov. sp. **Cole, Tennyson, Ksepka & Thomas**

Holotype

NMNZ S.33007 (AD161). Skull (see *Figure 10a1 – a3*). A DNA sequence was obtained from the cranium. A thoracic vertebra apparently associated with this skull is not considered to be part of the holotype. This specimen was originally identified as *Eudyptes* sp.

Paratypes

NMNZ S.24277 (AD291). Left carpometacarpus (see *Figure 10e1 – e2*). A DNA sequence was obtained from this specimen. Collected: B. G. McFadgen, Jan 1987, Waitangi West, Chatham

Island. This specimen was originally identified as *Eudyptes sclateri*. Age: past 7000 years cal BP (Millener, 1999).

NMNZ S.25157 (AD156). Right humerus (see *Figure 10e1 – e2*). A DNA sequence was obtained from this specimen. Collected: P. R. Millener, 10 Jan 1988, Maunganui, inland series of dune hollows, Chatham Island. This specimen was originally identified as *Eudyptes sclateri*. Age: past 7000 years cal BP (Millener, 1999).

NMNZ S.26908. Skull. DNA extraction was attempted but was unsuccessful. Collected: P. R. Millener, 21 Feb 1989, Maunganui, inland series of dune hollows, Chatham Island. This specimen was originally identified as *Eudyptes pachyrhynchus* and later figured by Millener (1999) as representing an undescribed *Eudyptes* species. Age: past 7000 years cal BP (Millener, 1999).

NMNZ S.27259 (AD158). Right coracoid (see *Figure 10d1 – d2*). Collected: H. O. Forbes, 1892, Chatham Islands. DNA sequence obtained. Originally identified as *Eudyptes* sp. with an associated cranium labelled *Catarrhactes schlegeli*. Age: past 7000 years cal BP (Millener, 1999).

NMNZ S.47917 (AD162). Left coracoid. A DNA sequence was obtained from this specimen. Collected: P. R. Millener, 19 Feb 1991, Waitangi West, Chatham Island. This specimen was originally identified as *Eudyptes*. Age: past 7000 years cal BP (Millener, 1999).

NMNZ S.30440. Mandible (see *Figure 10b*). DNA extraction was not attempted. Collected: P. R. Millener, 29 Jan 1992, Kaingaroa, foredune, Chatham Island. Figured by Millener (1999) as representing an undescribed *Eudyptes* species. Age: past 7000 years cal BP (Millener, 1999).

NMNZ S.47921 (AD352). Left tibiotarsus. A DNA sequence was obtained from this specimen. Collected: A. J. D. Tennyson, 1987 – 1988, Mangere Island, Chatham Islands. This specimen was originally identified as *Eudyptes* ?n. sp. Age: from a midden, possibly 19th century (Tennyson & Millener, 1994).

CM Av.6816 (AD157). Largely complete skull. A DNA sequence was obtained from this specimen. Collected: 1950, Chatham Is. This specimen was originally identified as *Eudyptes pachyrhynchus*. Age: past 7000 years cal BP (Millener, 1999).

CM Av.27407 (AD345). Left humerus. A DNA sequence was obtained from this specimen. Collected: R. J. Scarlett *et al.*, 27 Jan 1973, Wharekauri, Chatham Island. This specimen was

originally identified as *Eudyptes pachyrhynchus sclateri.* Age: past 7000 years cal BP (Millener, 1999).

CM Av.27867 (AD250). Left humerus. A DNA sequence was obtained from this specimen. Collected: R. J. Scarlett, 9 Dec 1972, Dunes, Te One, Chatham Island. This specimen was originally identified as *Eudyptes pachyrhynchus.* Age: past 7000 years cal BP (Millener, 1999).

Referred specimens

Sampling number AD282. Left humerus. A DNA sequence was obtained from this specimen. Collected: NMNZ MI/II/4-60 (rust + 5r), Centennial Inn, Paekakariki, Near Wellington. This specimen was originally identified as *Eudyptes pachyrhynchus*. Age: Early Māori from a midden deposit dated to 733 – 566 cal BP (Davidson, 1988).

Sampling number AD309. Right humerus. A DNA sequence was obtained from this specimen. Collected: NMNZ NRO350, Black Rocks, Black Midden, BR3 N168-9/77, 3/2/8. This specimen was originally identified as *Eudyptes pachyrhynchus*. Age: Early Māori from a midden deposit dated to 733 – 566 cal BP (Scarlett, 1979).

Sampling number AD391. Part of a coracoid. A DNA sequence was obtained from this specimen. Collected: NMNZ NRO344, Black Rocks, Cresent, BR4 Box 2a bird, 4/2/3. This specimen was originally identified as *Eudyptes pachyrhynchus*. Age: Early Māori from a midden deposit dated to 733 – 566 cal BP (Scarlett, 1979).

Sampling number AD394. Trochlea. A DNA sequence was obtained from this specimen. Collected: NMNZ NRO339, Black Rocks, Cresent, BR4 Box 2a 1 (bird) 4/1/6. This specimen was originally identified as *Eudyptes pachyrhynchus.* Age: Early Māori from a midden deposit dated to 733 – 566 cal BP (Scarlett, 1979).

CM Av.21751 (AD248). Part of a right radius. A DNA sequence was obtained from this specimen. Rakauteru Cave. 16 km north of Kaikoura. Collected: B.N. Norris, Nov/Dec 1967 – Jan 1968. This specimen was originally identified as *Eudyptes pachyrhynchus*. Age: Archaeological specimen with no radiocarbon date.

CM Av.25895 (AD342). Left tibiotarsus. A DNA sequence was obtained from this specimen. Shelter, Le Bons Bay, Banks Peninsula. Collected: R.S. Duff, 22 Oct 1955. This specimen was originally identified as *Eudyptes p. pachyrhynchus*. Age: Archaeological specimen with no radiocarbon date.

Etymology

The specific epithet honours John Warham (1919–2010), who carried out pioneering studies on *Eudyptes* penguins.

Type locality

Foredune about 200 m west of Tahatika Creek, Chatham Island. Collected by P. R. Millener, 20 Jan 1993. Age: less than 7000 years cal BP, which is the maximum age of the dunes. A Chatham Island duck (*Anas chathamica*) bone from the site dates to 1529 ± 57 ^{14}C BP (1405 – 1185 cal BP) (Millener, 1999).

Paratype and stratigraphic context

Most bones in the type series were isolated elements collected from eroded dune surfaces. However, the paratype from Mangere Island (NMNZ S.47921) was from a soil deposit that contained European-era remains (but had possible rabbit disturbance). No *E. warhami* bones were found in articulated association.

Diagnosis

E. warhami is characterized by an elongate ovoid premaxilla in dorsal view and a relatively shallow mandible. The species is distinguished from *E. chrysocome*, *E. filholi* and *E. moseleyi* by size (*Supplementary Table 7*). The largest *E. warhami* specimens (including the holotype) rival *E. chrysolophus schlegeli*, which is the largest extant *Eudyptes* taxon. *E. warhami* is distinguished from *E. pachyrhynchus* and *E. robustus* by a relatively elongate premaxilla, is distinguished from *E. chrysolophus chrysolophus* and *E. chrysolophus schlegeli* by a proportionally shallower mandible and is distinguished from *E. sclateri* by a more bowed premaxilla (dorsally) and a notably shallower mandible. *E. warhami* is distinguished from the Pliocene *E. calauina* by a smaller and more slender humerus (maximum length 70 mm in *E. warhami* versus ~81 mm in *E. calauina*).

Distribution

E. warhami was presumably once widespread along the coastlines of the Chatham Islands archipelago. The type series includes specimens from northern Chatham Island (43.71°S) to Mangere Island (44.27°S). The referred specimen series indicates that the species ranged westward to the east coast of mainland NZ (see also Cole *et al*., 2019a; *Chapter 4*).

Description

Eudyptes penguins exhibit sexual size dimorphism so inter-specific size variation would be expected in *E. warhami*. The beak shape varies subtly between most *Eudyptes* species, but post-cranial bones are largely inseparable based on shape. Molecular data circumscribing the *E. warhami* material confirms that the species fits this pattern. As is typical of larger *Eudyptes* species, the beak of *E. warhami* is more elongate than in smaller congeners. The rostral portion of the upper beak is markedly swollen as in *Eudyptes* and unlike *Megadyptes*. The jugal bar is strongly curved (see *Figure 10a1*), as in *Eudyptes*, and more so than in *Megadyptes*. The deep salt gland fossae are bounded by a shelf of bone. The mandibular ramus deepens strongly near the midpoint, a feature observed within *Eudyptes* and *Pygoscelis* penguins that is associated with a preference for planktonic prey (Zusi, 1975). Similar to extant *Eudyptes* species, the coracoidal fenestra is completely enclosed by a bridge of bone extending from the procoracoid process. The humerus is relatively wide, lacks a pronounced notch between the head and the dorsal tubercle, and has a posterior trochlear process which projects beyond the ventral border of the shaft. Skull length in *E. warhami* is comparable to skull length in the largest living *Eudyptes* penguins including *E. chrysolophus schlegeli* (*Supplementary Table 7*). However, coracoids, a humerus and a carpometacarpus from other *E. warhami* individuals are smaller than equivalent elements from *E. chrysolophus schlegeli* and instead are sized between *E. sclateri* and *E. robustus* (*Supplementary Table 7*). If sizes for cranial and postcranial elements represent within-population variation (e.g. sexual dimorphism) then body length of *E. warhami* was approximately 62 – 70 cm in comparison with other *Eudyptes* penguins (Stonehouse, 1967). The nomenclatural act for *E. warhami* has been registered to ZooBank (urn:lsid:zoobank.org:act:640EE978-13B5-4913-BB4E-6E16395325AC).

Comparison (measurements)

Bill length is 55% of the total skull length in *E. warhami* and ranges from 51% in *E. filholi* and *E. pachyrhynchus* to 59% in *E. chrysolophus schlegeli* (*Supplementary Table 7*). Caudal-most width of the bill is 29% of the bill length in *E. warhami* and similarly narrow in *E. chrysolophus schlegeli* (27%), but much broader in other extant *Eudyptes* species (e.g. 35% in *E. sclateri* and 41% in *E. filholi*). The maximum width of the internarial bar is 7% of the total skull length in *E. warhami* and 6% in other extant *Eudyptes*. Calvarium width is 86% of the calvarium length in *E. warhami* and 89 – 91% in other extant *Eudyptes* penguins. The mandible depth is 17% of the mandible length in *E. warhami* and is typical of extant *Eudyptes* penguins (e.g. 17% in *E. robustus*; 18% in *E. moseleyi*). The mandible is relatively deep, accounting for 21% of the mandible length in *E. chrysolophus chrysolophus* and *E. sclateri*.

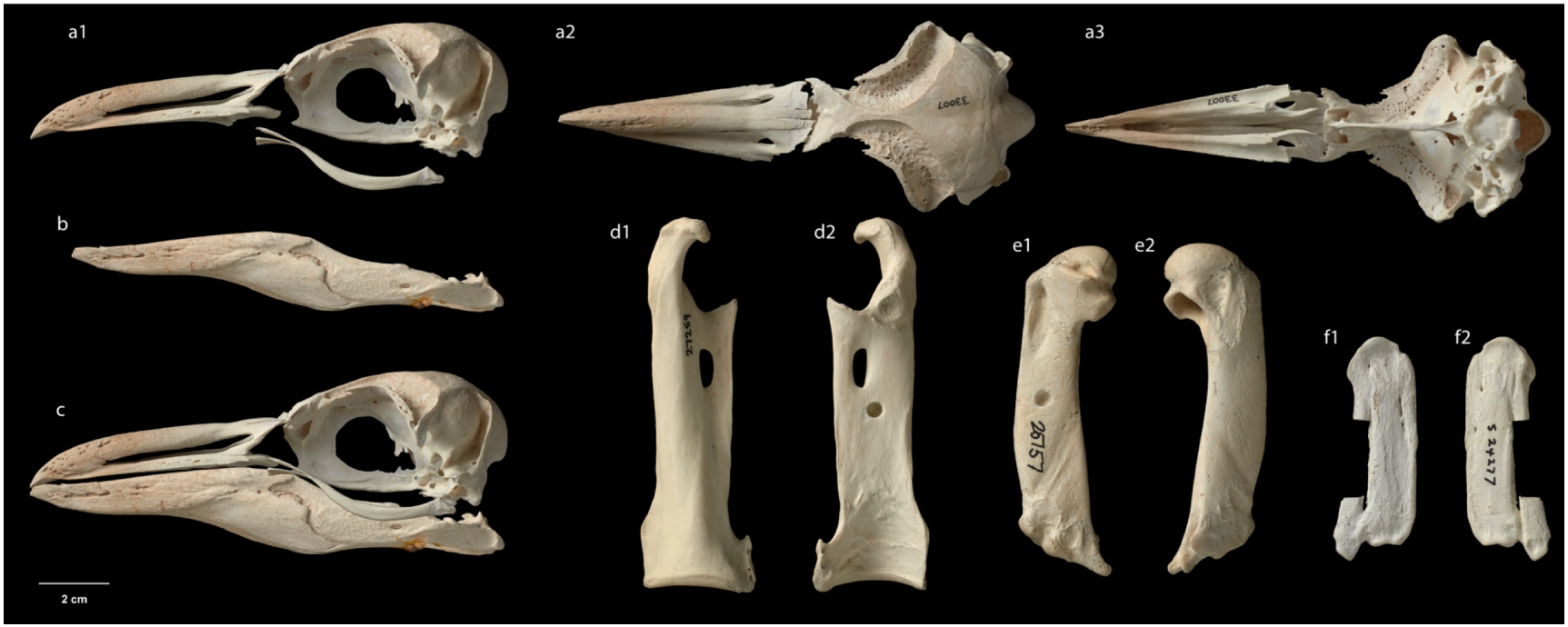


Figure 10. *Eudyptes warhami* holotype and paratypes. **(a)** holotype cranium NMNZ S.33007 in (a1) lateral, (a2) dorsal, and (a3) ventral views; **(b)** paratype mandible NMNZ S.25157 in lateral view; **(c)** composite skull digitally reconstructed by combining holotype cranium and re-scaled paratype mandible to illustrate beak depth; **(d)** paratype right coracoid NMNZ S.27259 in (d1) ventral and (d2) dorsal views, (e) paratype right humerus in (e1) cranial) and (e2) caudal views; **(e)** paratype carpometacarpus NMNZ S.24277 in (e1) ventral) and (e2) dorsal views. DNA sequences were obtained from all specimens except for the mandible. Photos: Jean-Claude Stahl, NMNZ.

Megadyptes Milne-Edwards, 1880

Megadyptes antipodes (Hombron & Jacquinot, 1841)

Megadyptes antipodes richdalei nov. ssp., **Tennyson & Cole**

Holotype

NMNZ S.26921 (AD95). Partial skeleton (see *Figure 11b – p*). A DNA sequence was obtained from a rib. This specimen was originally identified as *Megadyptes antipodes*.

Paratypes

NMNZ S.47918 (AD159). Right coracoid. A DNA sequence was obtained from this specimen. Collected: P.R. Millener, 19 Feb 1991, Waitangi West, Chatham Island. This specimen was originally identified as *Eudyptes*. Age: past 7000 years cal BP (Millener, 1999).

NMNZ S.30968. Mandible. DNA extraction was not attempted. Collected: P. R. Millener, 2 Feb 1992; Maunganui Beach – c.500m west of Takehanga Stream "parking spot", Chatham Island. This specimen was originally identified as *Eudyptes*, re-identified on 5 Apr 2016 as *Megadyptes*. Age: past 7000 years cal BP (Millener, 1999).

NMNZ S.47765. Premaxilla. DNA extraction was not attempted. Collected: P. R. Millener & N. H. Hyde, 12 Feb 1992; Kaingaroa, hill site, Chatham Island. This specimen was originally identified as *Eudyptes*, re-identified as *Megadyptes* based on premaxilla shape. Age: past 7000 years cal BP (Millener, 1999).

NMNZ S.45876 (AD88). Skull (*Figure 11a1 – a3*). A DNA sequence was obtained from this specimen. Collected: H. O. Forbes, Chatham Islands. This specimen was originally identified as *Megadyptes*. Age: past 7000 years cal BP (Millener, 1999).

CM Av11287 (AD252). Left humerus. A DNA sequence was obtained from this specimen. Collected: J. Eyles, 1952, “Chatham Is”. This specimen was originally identified as *Eudyptes pachyrhynchus*. Age: past 7000 years cal BP (Millener, 1999).

CM (ACAD12997) Unregistered Canterbury Museum. Left humerus (see *Supplementary Figure 13*). A DNA sequence was obtained from this specimen. Collected: R. P. Scofield, Te One Dunes, Chatham Island, 2007. This specimen was originally identified as *Eudyptes*. Age: past 7000 years cal BP (Millener, 1999).

AM LB12063 (AD301). Part left tibiotarsus. A DNA sequence was obtained from this specimen. Collected: M. K. Eagle, northwest Pitt Island, 2002. This specimen was originally identified as *Eudyptes* sp. Age: past 7000 years cal BP (Millener, 1999).

Etymology

The subspecies epithet honours Lance Richdale (1900–1983), who carried out pioneering studies on *Megadyptes* ecology.

Type locality

Foredune east of Maunganui, Chatham Is. Collected by P. R. Millener, 21 Feb 1989. Age: <7000 years cal BP, which is the maximum age of the dunes. A Chatham Island rail (*Diaphorapteryx hawkinsi*) bone from the site dates to 1860 ± 150 ^{14}C BP (1996 – 1314 cal BP) (Millener, 1999).

Paratype and stratigraphic context

All bones in the type series were collected from eroded dune surfaces and were from the Chatham Islands archipelago.

Diagnosis

A *Megadyptes* penguin, smaller than *M. antipodes* and *M. waitaha* (*Supplementary Tables 8 – 12*). *M. a. richdalei* nov. ssp. represents a genetic lineage comprising distinct haplotypes (>15 private mitochondrial SNPs) which were not detected in other living or extinct *Megadyptes* populations (*Figure 8* and *Supplementary Tables 5 – 6*).

Distribution

Chatham and Pitt Islands (from 43.73°S to 44.23°S). Presumably *M. a. richdalei* occurred around all coasts of the Chatham Islands archipelago.

Description

M. a. richdalei is the smallest *Megadyptes* penguin, being on average 10 – 15% smaller than *M. antipodes,* and approximately 5% smaller than *M. waitaha* (though size distributions overlap; see *Supplementary Table 8*). Both recently-extinct taxa are smaller than *M. antipodes*, with almost no overlap between the extant and extinct taxa in bone lengths. The skull closely resembles that of *M. antipodes* and differs from the contemporaneous *E. warhami* in its more slender upper beak and shallower mandible (without a pronounced deepening at the midpoint). We observed no osteological differentiation between the three *Megadyptes* taxa that could not

be accounted for either by size or individual variation (as reflected in *M. antipodes* specimens), suggesting the proposed postcranial differences between *M. antipodes* and *M. waitaha* (Boessenkool *et al*., 2008) cannot consistently differentiate these taxa. Thus, we consider the Chatham *Megadyptes* taxon to represent an instance of isometric dwarfing and recommend the two recently extinct forms be recognized as subspecies of *Megadyptes antipodes.* We follow the NZ Bird Checklist Committee (Gill *et al*., 2010) in defining subspecies using the Diagnostic Species Concept, where it is expected that a subspecies will meet 75% diagnosable criteria (Amadon, 1949; Patten & Unitt, 2002). The size uncertainty for *M. a. richdalei* arises from minor size variation recorded in appendicular elements from all three subspecies (*Supplementary Tables 8 – 12*). Elements in the cranial and appendicular skeletons generally scale isometrically across all three subspecies. The body length of *M. a. richdalei* is approximately 55 – 60 cm (Stonehouse, 1967). The nomenclatural act for *M. a. richdalei* is registered in ZooBank (urn:lsid:zoobank.org:act:169A6F00-1564-4977-B56B-09005591A88E).

Comparison (measurements)

The maximum width of the internarial bar is 5% of the total skull length in both *M. a. richdalei* and *M. a. antipodes*. The mandible depth is 12% of the mandible length in *M. a. richdalei* and 11 – 13% in *M. a. antipodes*. The proximal width of the coracoid is 29 – 31% of the total coracoid length for *M. a. richdalei*, 30 – 3% for *M. a. antipodes* and 29 – 33% for *M. a. waitaha*. The maximum width of the humeral head (the humerus proximal width in *Supplementary Tables 8 – 12*) is 29 – 30% of the total length of the humerus in *M. a. richdalei* and *M. a. waitaha* and 28 – 31% in *M. a. antipodes*. Likewise, the maximum width at the distal end of the humerus is 28 – 31% of the total humerus length in *M. a. richdalei*, 26 – 29% in *M. a. waitaha* and 27 – 29% in *M. a. antipodes*. The total length of the humerus is 85% of the total length of the coracoid in *M. a. richdalei,* 84 – 87% in *M. a. antipodes* and 88 – 89% in *M. a. waitaha*. The total radius length is 76% of the total humerus length and 98% of the total ulna length in *M. a. richdalei*. The radius length ranges from 75 – 79% of the humerus length and from 96 – 99% of the humerus length in *M. a. antipodes*. The carpometacarpus length is 60% of the humerus length in *M. a. richdalei* and 57 – 60% of the length in *M. a. richdalei*. The width of the femur at midway along the shaft is 11% of the total length of the femur in *M. a. richdalei* and 10 – 11% in both *M. a. antipodes* and *M. a. waitaha*. The total length of the femur in *M. a. richdalei* is 106% of the length of the humerus; in *M. a. antipodes* this value is 107 – 112% and in *M. a. waitaha* this value is 108 – 110%. The width of the distal end of the tibiotarsus is 13% of the total length of the tibiotarsus in *M. a. richdalei* and 12 – 13% in *M. a.*

antipodes. The total length of the tibiotarsus in *M. a. richdalei* is 158% of the length of the femur and is 153 – 156% in *M. a. antipodes*. The maximum distal width of the tarsometatarsus was measured from the external surfaces of trochleae two and four, and expressed as a ratio of the total length of the tarsometatarsus from the distal surface of trochlea three to the proximal articular surface. The distal tarsometatarsus proportion for *M. a. richdalei* is 67% compared to 63 – 68% for *M. a. antipodes* and 58 – 69% for *M. a. waitaha*. The total length of the tarsometatarsus in *M. a. richdalei* is 43% of the length of the femur; in *M. a. antipodes* this value is 41 – 43% and in *M. a. waitaha* this value is 41 – 45%.

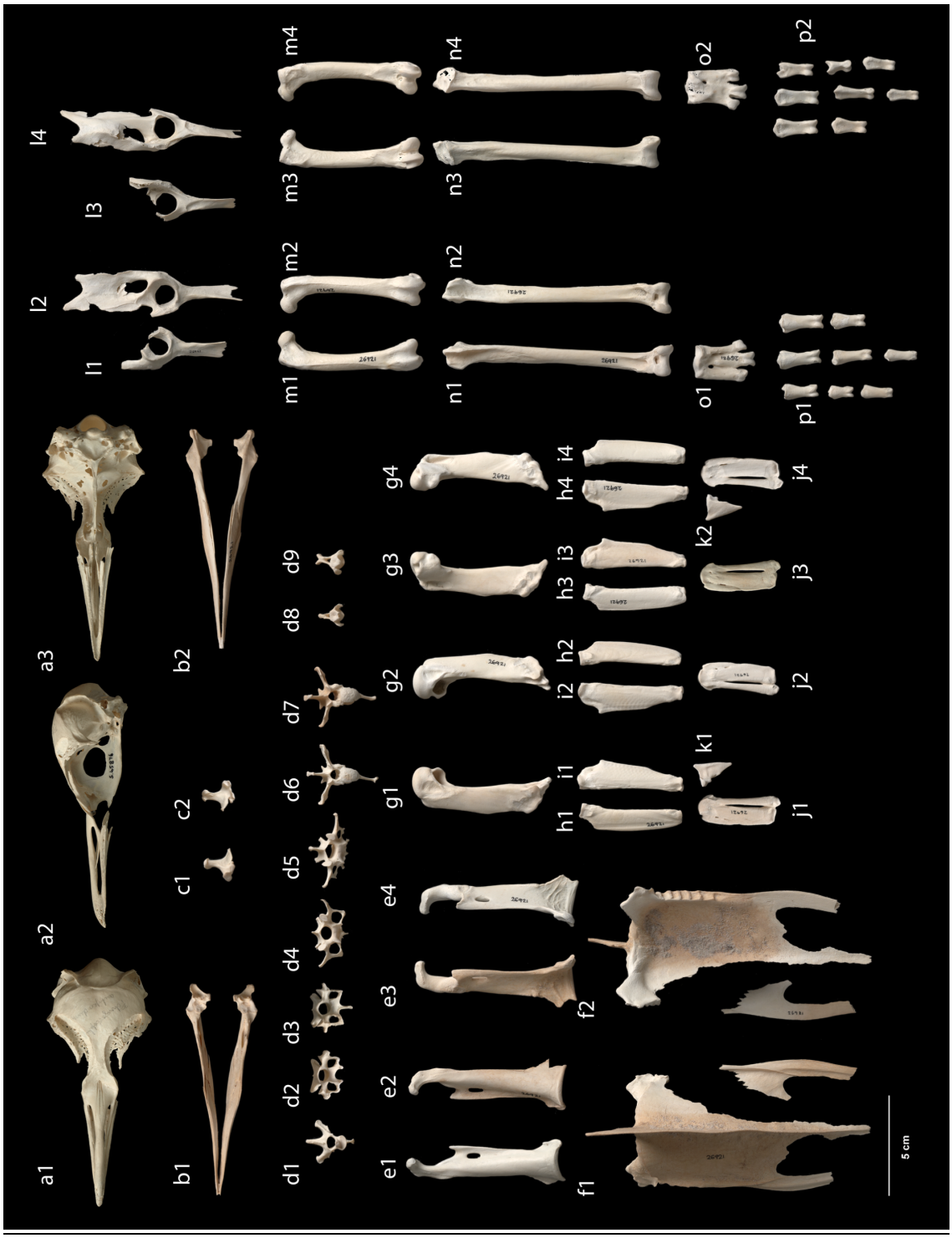


Figure 11 (overleaf)**.** *Megadyptes antipodes richdalei* holotype and paratypes. **(a)** Paratype skull NMNZ S.45876 in (a1) dorsal, (a2) lateral, and (a3) ventral views. **(b – p)** Holotype NMNZ S.26921 cranial and postcranial elements: (b) mandible in (b1) dorsal and (b2) ventral views, (c) left quadrate in (c1) lateral and (c2) medial views, (d1) axis, (d2) eight cervical vertebra, (d3) tenth cervical vertebra, (d4) twelfth cervical vertebra, (d5) thirteenth cervical vertebra, (d6+d7) undetermined caudal thoracic vertebrae, (d8+d9) undetermined caudal vertebrae, all in cranial view, (e1) right and (e2) left coracoid in ventral view, (e3) left and (e4) right coracoid in dorsal view, (f) sternum in (f1) ventral and (f2) dorsal view, (g1) left and (g2) right humerus in caudal view, (g3) left and (g4) right humerus in cranial view, (h1) left and (h2) right radius in dorsal view, (h3) left and (h4) right radius in ventral view, (i1) left and (i2) right ulna in dorsal view, (i3) left and (i4) right ulna in ventral view, (k) left ulnar in (k1) dorsal and (k2) ventral view, (j1) left and (j2) right carpometacarpus in dorsal view, (j3) right and (j4) left carpometacarpus in ventral view, (h3) left and (h4) right carpometacarpus in ventral view, acetabular region of the (l1) left and (l2) right side of the pelvis in lateral view, acetabular region of the (l3) left and (l4) right side of the pelvis in medial view, (m1) right and (m2) left femur in cranial view, (m3) left and (m4) right femur in caudal view, (n1) right and (n2) left tibiotarsus in cranial view, (n3) left and (n4) right tibiotarsus in caudal view, (o) right tarsometatarsus in (o1) cranial view and (o2) caudal view), (p) pedal phalanges in (p1) dorsal view and (p2 (ventral view). DNA sequences were obtained from both the holotype and paratype. Photos: Jean-Claude Stahl, NMNZ.

Details of Chatham Island *Eudyptes sclateri* fossil bones

NMNZ S.47919 (AD207). Right coracoid. A DNA sequence was obtained from this specimen. Collected: P. R. Millener, 19 Feb 1991, Waitangi West, Chatham Island. This specimen was originally identified as *Eudyptes*. Age: past 7000 years cal BP (Millener, 1999).

NMNZ S.37744 (AD297). Part left tibiotarsus. A DNA sequence was obtained from this specimen. Collected: A. J. D. Tennyson, 6 Dec 1997, on south bank of stream mouth 300 m south of Sandy Point, North Head, Pitt Island, Chatham Islands. This specimen was originally identified as *Eudyptes sclateri*. Age: past 7000 years cal BP (Millener, 1999).

CM AV7654 (AD86). Cranium. A DNA sequence was obtained from this specimen. Collected: Chatham Is., 1950, ?Kinsey collection. This specimen was originally identified as *Eudyptes* ?*pachyrhynchus*. Age: past 7000 years cal BP (Millener, 1999).

Discussion

Timing of penguin evolution is linked to island emergence

While some studies have suggested that crown penguins began radiating in the Eocene (Baker *et al*., 2006), our divergence estimates (*Figure 9C* and *Supplementary Table 13*) indicate a Miocene origin of crown penguins (Subramanian *et al*., 2013; Gavryushkina *et al*., 2017). Moreover, our results support the hypothesis that a large proportion of penguins diverged within the past two million years (Gavryushkina *et al*., 2017). We propose that this diversification pulse was tied to the emergence of islands, which created new opportunities for isolation and speciation.

While island emergence can spawn diverse biological radiations, several studies have detected island-endemic lineages substantially pre-dating island formation (Lewin, 1985; McCulloch & Waters, in press). McCulloch and Waters (in press) suggested that uncertainties in molecular dating, especially when divergences are ancient and sequence data is limited, can be problematic for branch-length estimation of deeply diverged lineages. However, our estimated divergence dates (based on near-complete mitogenomes and well-justified fossil calibration points) for most island-endemic penguin taxa (e.g. those restricted to one island/archipelago) are consistently younger than the geological estimates of the islands they inhabit (see *Figure 9C* and *Supplementary Table 13*). While the ancestral distributions of clades are not always clear, the finding of numerous recently-evolved taxa endemic to geologically-young islands (Fleischer *et al*., 1998; Mendelson & Shaw, 2005) strongly suggests that these lineages evolved *in situ*, rather than being relicts of formerly widespread species. Our study thus provides temporal genetic evidence linking penguin speciation to island formation (*Figure 9C*).

Our analysis found that *Eudyptes sclateri*/*E. warhami* probably diverged from *E. robustus*/*E. pachyrhynchus* 3.5 – 1.7 mya, following the emergence of the Chatham Islands (3 mya) (Campbell *et al*., 2008). Divergence estimates (95% HPD) for both *Megadyptes antipodes richdalei* (0.4 – 0.1 mya) and *Eudyptes warhami* (2.5 – 1.1 mya) similarly post-date the emergence of the Chatham Islands, and are concordant with divergence estimates for many endemic Chatham Islands lineages, including plants (Heenan *et al*., 2010), insects (Trewick, 2000) and birds (Mitchell *et al*., 2014a; Wood *et al*., 2014; see *Figure 12* and McCulloch & Waters, in press). While the geological age of The Snares uplift remains unclear (Mortimer *et al*., 2015), our analysis of *E. robustus* suggests that the islands have been emergent for at least 1.4 – 0.5 Ma (Nick Mortimer, personal communication). The divergence between *E. moseleyi* and *E. chrysocome*/*E. filholi* 2.7 – 1.2 mya corresponds with the emergence of Gough Island around 2.5 mya (Maund *et al*., 1988), and populations presumably dispersed to the younger islands of the Tristan da Cunha archipelago, Amsterdam Island, and St Paul Islands (McDougall & Ollier, 1982) via the ACC. The divergence between *E. chrysolophus chrysolophus* and the Macquarie Island-endemic *E. chrysolophus schlegeli* (0.2 – 0.0 mya) is concordant with the geological uplift of Macquarie Island 0.7 mya (Adamson *et al*., 1996). The Galápagos Island-endemic *Spheniscus mendiculus* diverged from its sister taxon *S. humboldti* 1.6 – 0.6 mya, shortly after the formation of several islands within this young archipelago (2 mya). Similar founder speciation has previously been inferred for numerous Galápagos Island-endemic taxa (Parent *et al*., 2008), including invertebrates (Parent & Crespi, 2006; Sequeira *et al*., 2008), reptiles (Caccone *et al*., 1999) and birds (Bollmer *et al*., 2006).

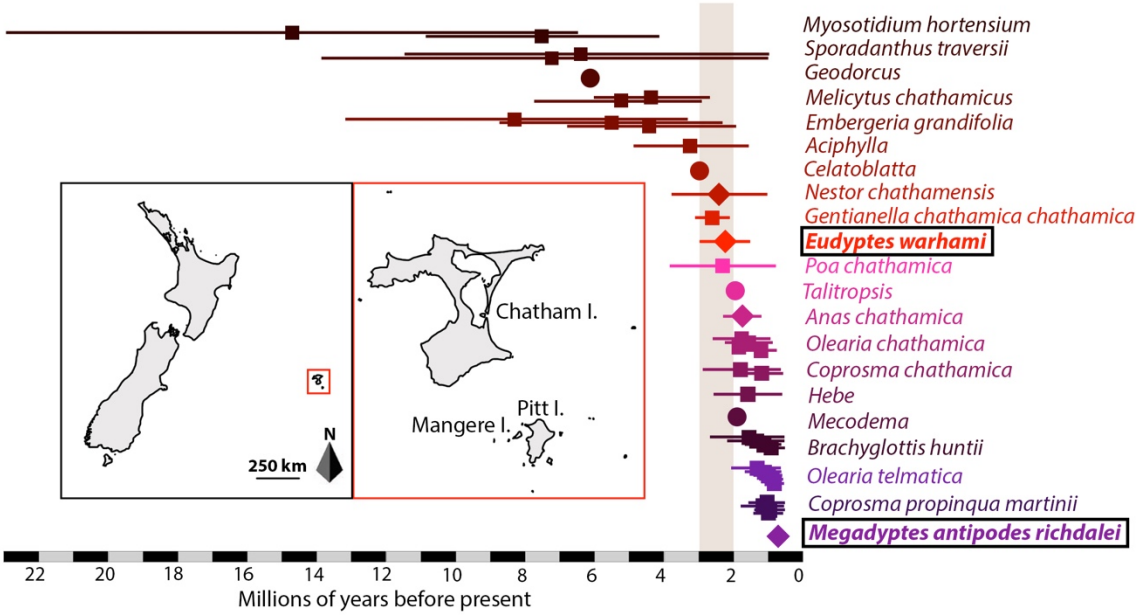


Figure 12. Estimated divergence dates of endemic Chatham Islands taxa. The divergence date of each taxon has been estimated from its nearest non-Chatham Islands relative, comprising plants (squares), invertebrates (circles) and birds (diamonds). Ages of plants are taken from Heenan *et al.* (2010). Some plant taxa were analysed using multiple methods, so for those taxa, each analysis is shown together in a cluster. Invertebrates are taken from Trewick (2000), and for each genus the divergence was calculated using 2% sequence divergence rate/Ma. Birds are taken from Mitchell *et al.* (2014a) and Wood *et al*. (2014) and are based on fossil calibrations. The 95% HPD for *Eudyptes warhami* and *Megadyptes antipodes richdalei* are shown. The thick brown bar represents the emergence of the Chatham Islands. Inset shows NZ and the Chatham Islands.

Our finding that many recent speciation events among penguins are temporally linked to island formation may provide important clues for understanding evolutionary patterns in other island-endemic taxa. Islands are clearly speciation hotspots for terrestrial taxa, but the role of island-emergence as a driver of speciation in marine taxa remains less clear (however see Mitchell *et al*. 2014a). The shag genus *Leucocarbo* similarly has endemic taxa associated with almost every sub-Antarctic island (Marchant & Higgins, 1990), providing a possible parallel example of recent founder speciation in the Southern Ocean. Time-calibrated genomic analysis provides an exceptional new tool for understanding the origins of such iconic Southern Ocean biodiversity.

Vulnerability of island taxa to human-induced extinctions

Our study uncovered two new island-endemic penguin taxa: *Eudyptes warhami* and *Megadyptes antipodes richdalei*. The presence of their bones in middens, and lack of reliable historical sightings, suggests that these taxa were extirpated shortly after human settlement on the Chatham Islands post 13th century CE (Maxwell & Smith, 2015). These findings thus potentially represent important new examples of human-driven, Holocene extinctions in the Pacific. *Eudyptes warhami* bones (*cf. E.* clade X; Cole *et al*., 2019a; *Chapter 4*) excavated from

coastal middens demonstrate that the species was also hunted on mainland NZ (see Cole *et al*., 2019a; *Chapter 4*). However, this does not prove the presence of a local breeding colony. In fact, many extant island-endemic *Eudyptes* species disperse widely during the non-breeding period: *E. pachyrhynchus* (breeds only on mainland NZ) and *E. robustus* (breeds only on The Snares) are commonly observed in southern Australia during winter (Woehler, 1992; Cole *et al*., 2018; *Chapter 5*; Mattern *et al*., 2018); and *E. sclateri* (breeds only on the Antipodes Islands) is commonly observed in NZ (Robertson *et al*., 2017). As *E. warhami* is relatively rare in mainland subfossil penguin assemblages (represented by only seven specimens among hundreds of genetically-identified penguin bones; Boessenkool *et al*., 2008; Rawlence *et al*., 2015a; Grosser *et al*., 2016; Cole *et al*., 2019a; *Chapter 4*), these mainland records fit the pattern that would be expected for non-breeding *E. warhami* individuals. In contrast, no *Megadyptes antipodes richdalei* bones have been detected in mainland subfossil assemblages (Boessenkool *et al*., 2008; Rawlence *et al*., 2015a; Cole *et al*., 2019a; *Chapter 4*). This pattern fits with the limited dispersal exhibited by extant *Megadyptes* populations (Boessenkool *et al*., 2009).

Conclusions

We found strong evidence for a Neogene radiation of crown penguins, and provide the first compelling evidence that island emergence drove Plio-Pleistocene penguin diversification. Such processes may also have driven diversification in the deeper past, as fossil data show much higher penguin diversities than present once existed in NZ (Ksepka & Ando, 2011), Antarctica (Jadwiszczak, 2006), Australia (Park & Fitzgerald, 2012), and Africa (Thomas & Ksepka, 2013). However, as most fossils from these regions are restricted to continental localities, and many islands have scant fossil records, the role of island formation in penguin diversification in the deep past remains obscured. Accordingly, if rates of island-mediated speciation were as high throughout the Cenozoic as in the Plio-Pleistocene, it is conceivable that fossils for a major proportion of extinct penguin taxa will never be found.

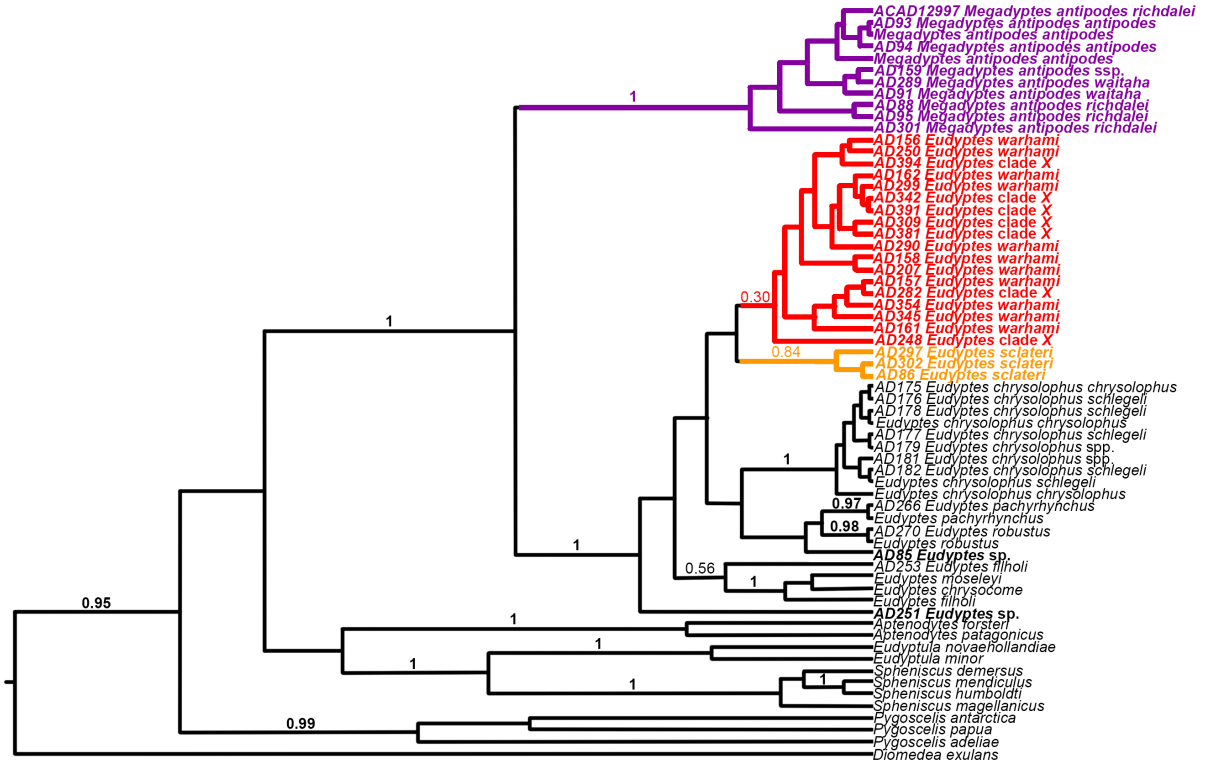
Previous studies based on traditional species concepts have struggled to account for recently-evolved biological diversity. Particularly relevant are scenarios of species 'divergence with geneflow', where introgression may occur among closely-related linages (Rheindt & Edwards, 2011). Although introgression between closely related penguin species within *Spheniscus* and *Eudyptes* has occasionally been reported (Napier, 1968; Woehler & Gilbert, 1990; White & Clausen, 2002; Simeone *et al*., 2009; Morrison & Sagar, 2014) solid confirmatory genetic evidence is lacking (see also *Chapter 3*). While this current study does not address the

possibility of introgression among penguin taxa (see Mays *et al*., in press), genome-wide SNP analyses provide insights into this question for penguins (*Chapter 3*).

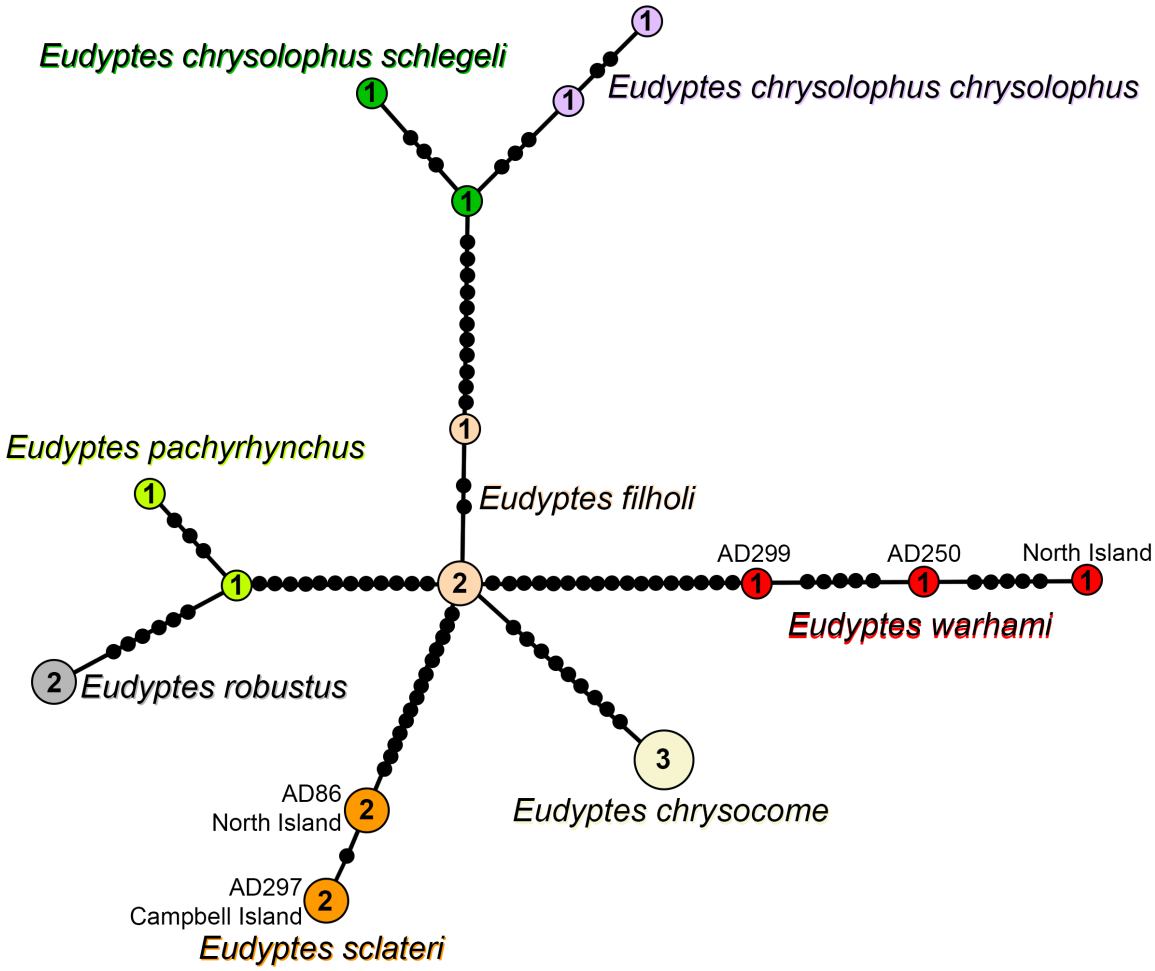
While our results reinforce the importance of islands in generating biodiversity, they also underscore the role of humans as agents of biodiversity loss, especially via the extinction of island-endemic taxa (Duncan *et al*., 2013; Valente *et al*., 2019). Today only *Eudyptula minor* breeds on the Chatham Islands, yet 500 years ago the archipelago held substantial penguin diversity, with two endemic taxa (*Eudyptes warhami* and *Megadyptes antipodes richdalei*; see *Figure 13*) alongside *Eudyptula minor* and possibly *Eudyptes sclateri*. As many of the bones were from middens, our results provide direct evidence that *E. warhami* was hunted by humans. Although no *Megadyptes antipodes richdalei* remains examined in this study were directly associated with human activity, the near-simultaneous disappearance of both this subspecies and *Eudyptes warhami* suggests that both extirpations were linked to the arrival of humans to the Chatham Islands. Our results further emphasise the value of aDNA for elucidating biodiversity shifts, including the dramatic rise and fall of island avifauna (Waters & Grosser, 2016).

Figure 13. Artist impression of *Eudyptes warhami* and *Megadyptes antipodes richdalei*. *Eudyptes warhami* is in the foreground and a single *Megadyptes antipodes richdalei* is in the background (far right). Artist: Sean Murtha, 2018.

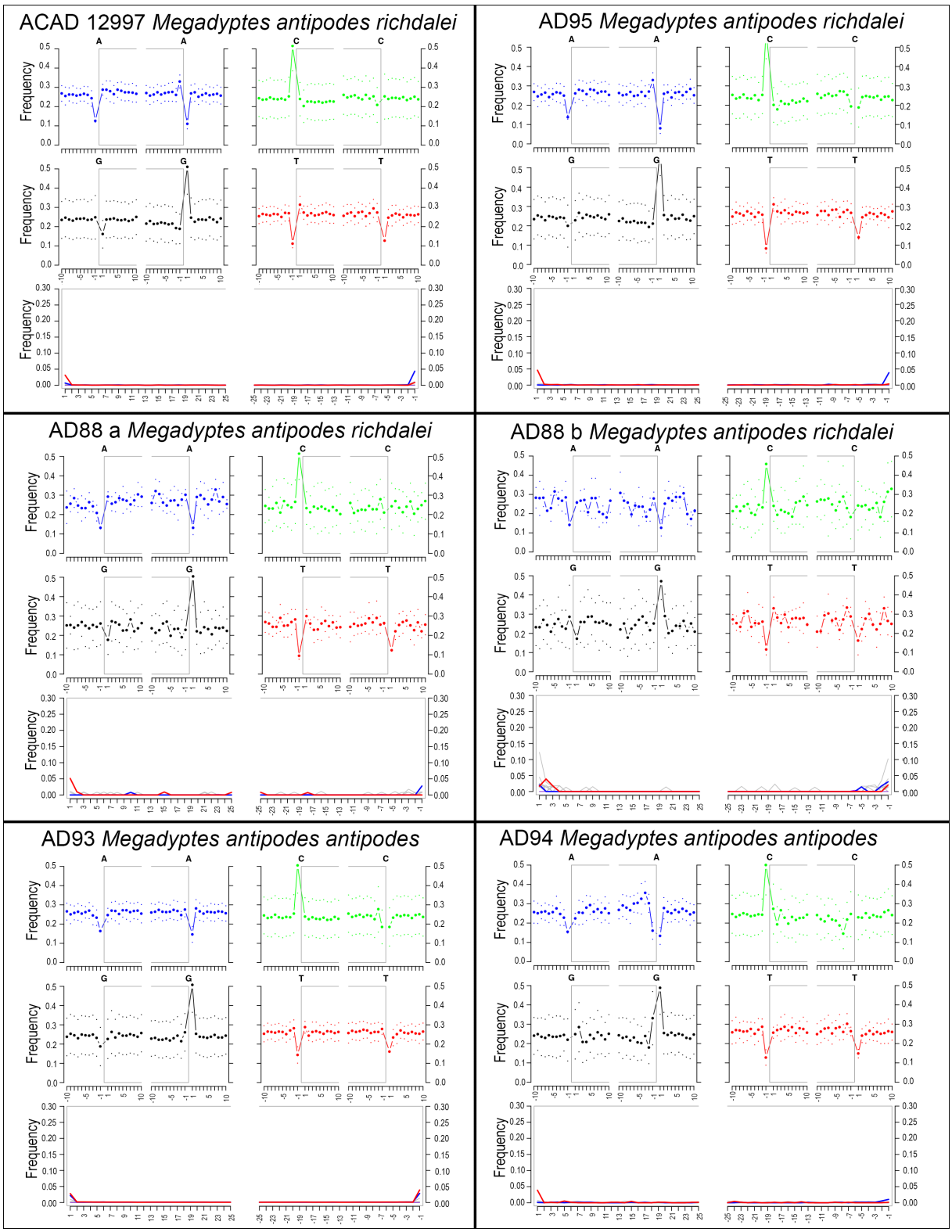
Supplementary Figures for *Chapter 2*



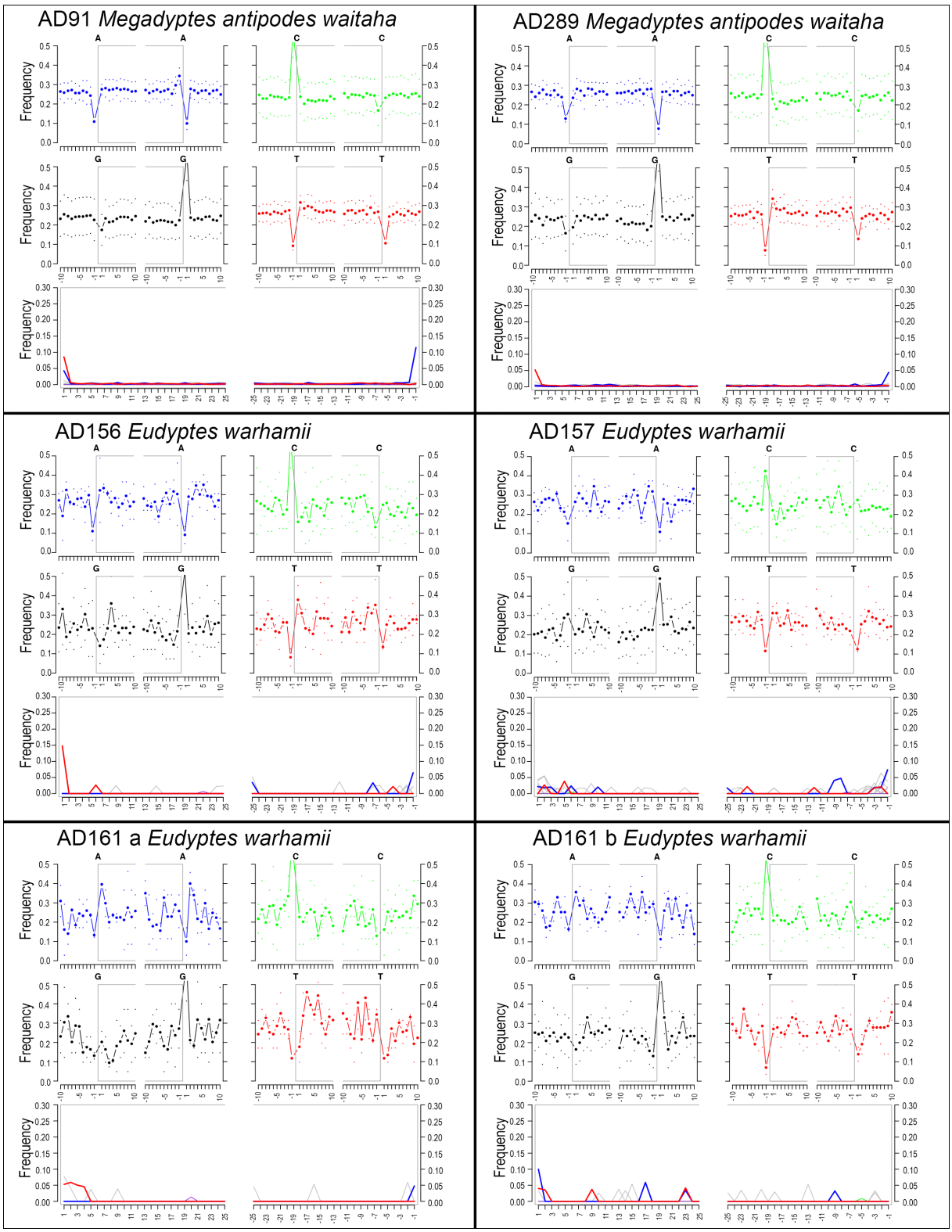
***Supplementary Figure 1*.** Bayesian phylogenetic tree of penguins, with all aDNA COI sequences from Chatham Islands bone samples. Prehistoric samples have AD numbers. Four bones from the Chatham Islands are identified as *Megadyptes antipodes richdalei*, one is *Megadyptes antipodes* ssp. (*Megadyptes* in purple), two are *Eudyptes sclateri* (orange) and 11 fall with *Eudyptes* clade X from mainland NZ (red; discussed in Cole *et al*., 2019a; *Chapter 4*]). Posterior probabilities >0.95 for the main clades are shown, and two lower posterior probabilities for *Eudyptes filholi* and *Eudyptes* sp. are shown.



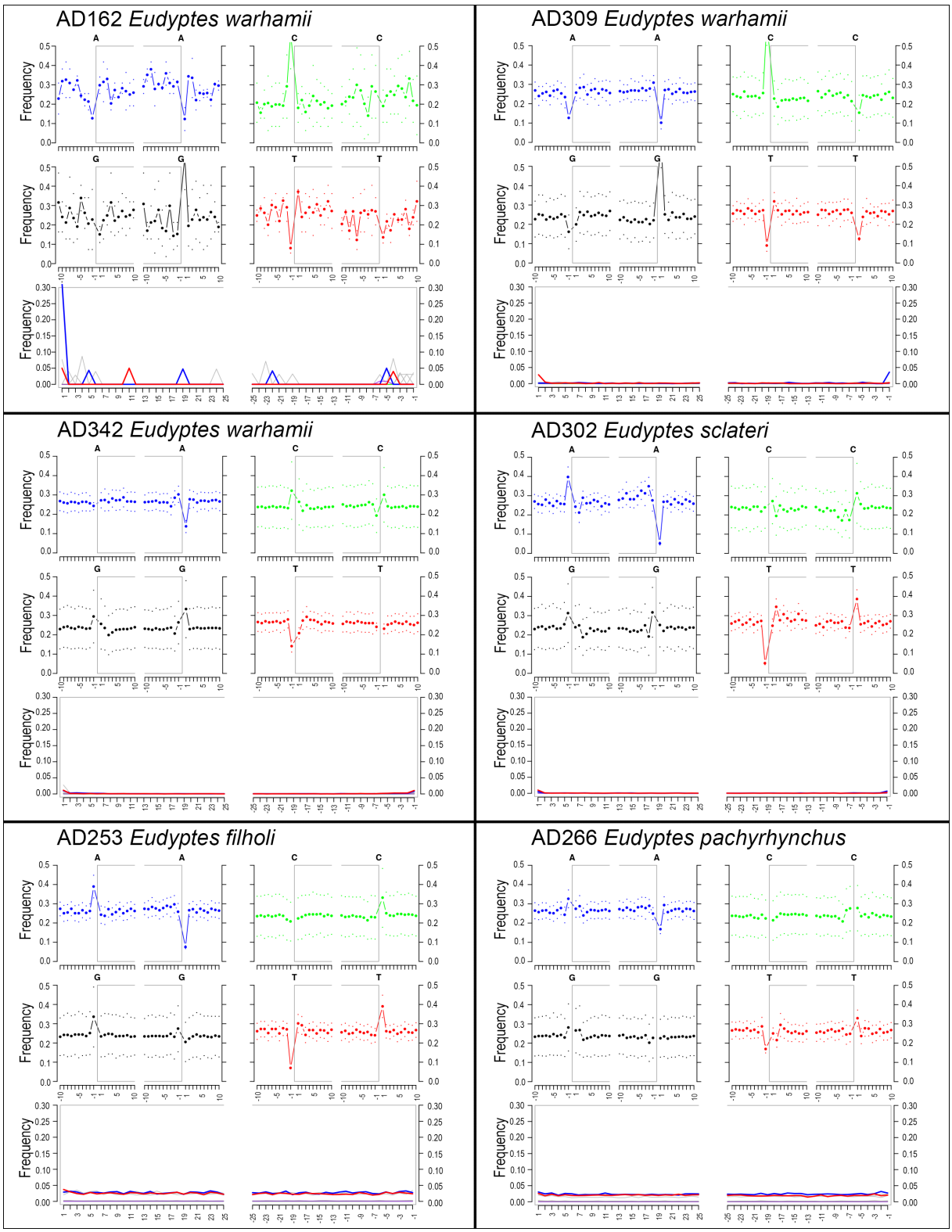
Supplementary Figure 2. Minimal-spanning haplotype network for *Eudyptes* penguins based on aDNA CR sequences. Coloured circles represent different taxa, with the number within and size of the circle representing the number of samples that had that haplotype. Small black circles represent haplotypes that were absent from the dataset. *E. warhami* (*cf*. *E*. clade X) samples are in red. Chatham Islands *E. sclateri* are shown in orange (AD86 and AD297) compared to *E. sclateri* from the NZ North Island and Campbell Island. Note, AD297 (Chatham Islands) shares a haplotype with an *E. sclateri* sample from Campbell Island, and AD86 (Chatham Islands) shares a haplotype with an *E. sclateri* sample from the North Island.



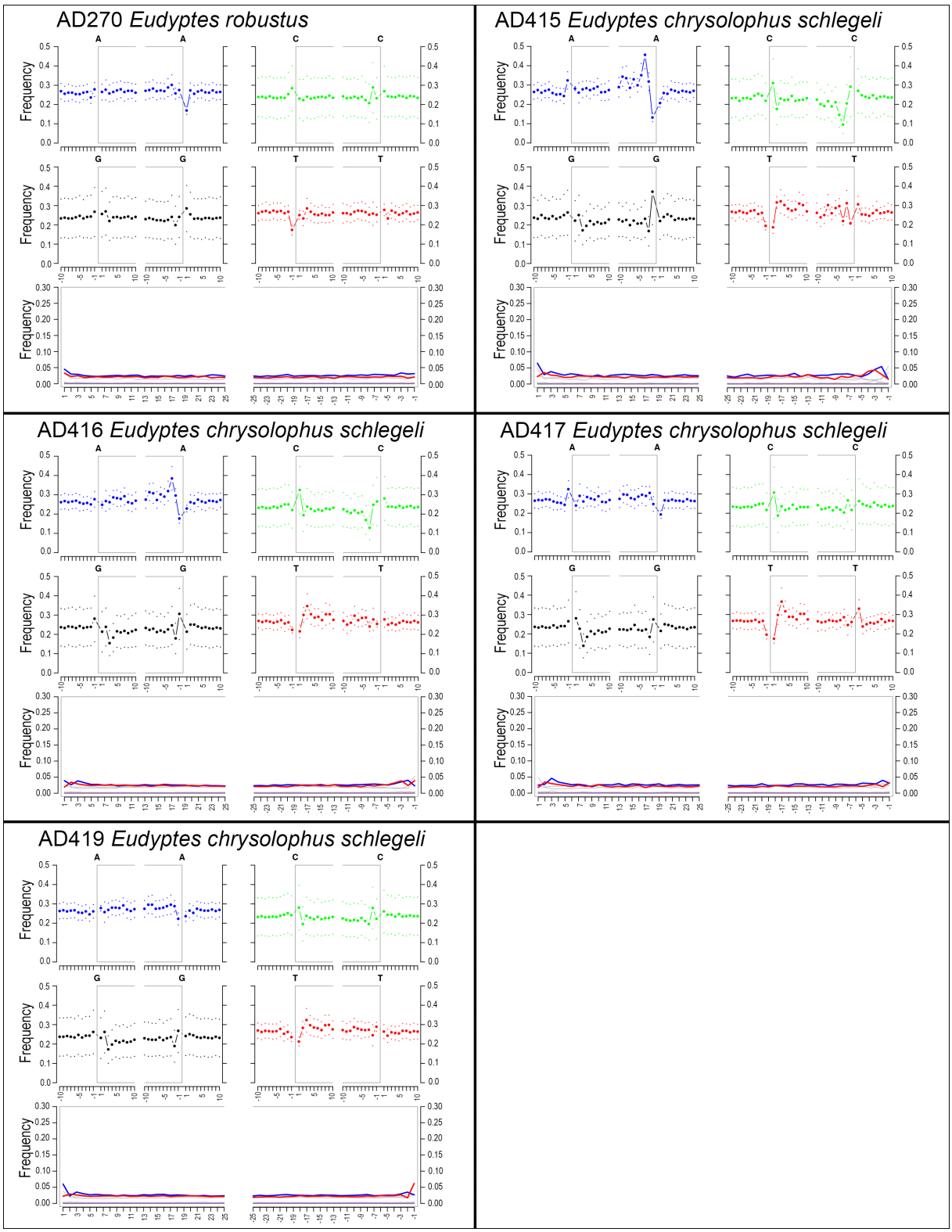
Supplementary Figure 3. MapDamage reports for the final round of mapping of six partial ancient mitogenomes to the *Megadyptes antipodes richdalei* (ACAD 12997) mitogenome. The top panels show the characteristic high frequency of purines (A and G) immediately before the reads. The lower panels show the accumulation of 5' C to T (red) and 3' G to A (blue) misincorporations characteristic of aDNA.



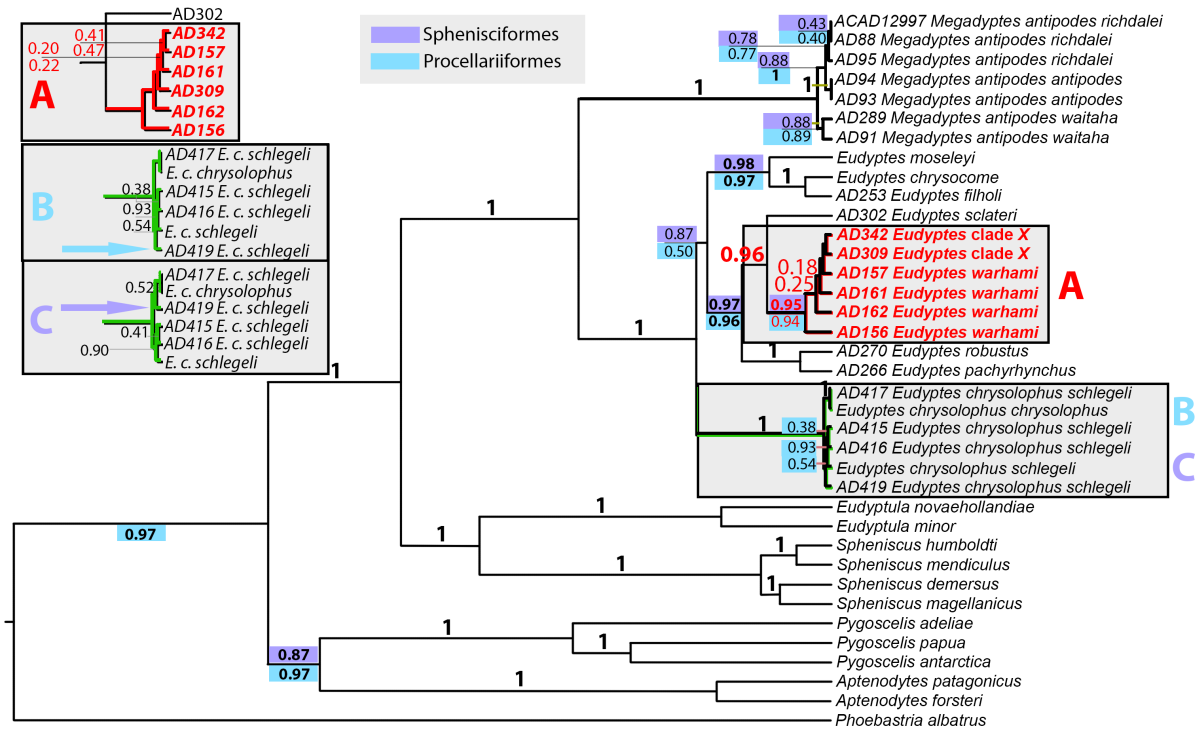
Supplementary Figure 4. MapDamage reports for the final round of mapping of six partial ancient mitogenomes to the *Megadyptes antipodes richdalei* (ACAD 12997) mitogenome (for all *Megadyptes* samples) or to the *Eudyptes sclateri* mitogenome (for all *Eudyptes* samples). The top panels show the characteristic high frequency of purines (A and G) immediately before the reads. The lower panels show the accumulation of 5' C to T (red) and 3' G to A (blue) misincorporations characteristic of aDNA.



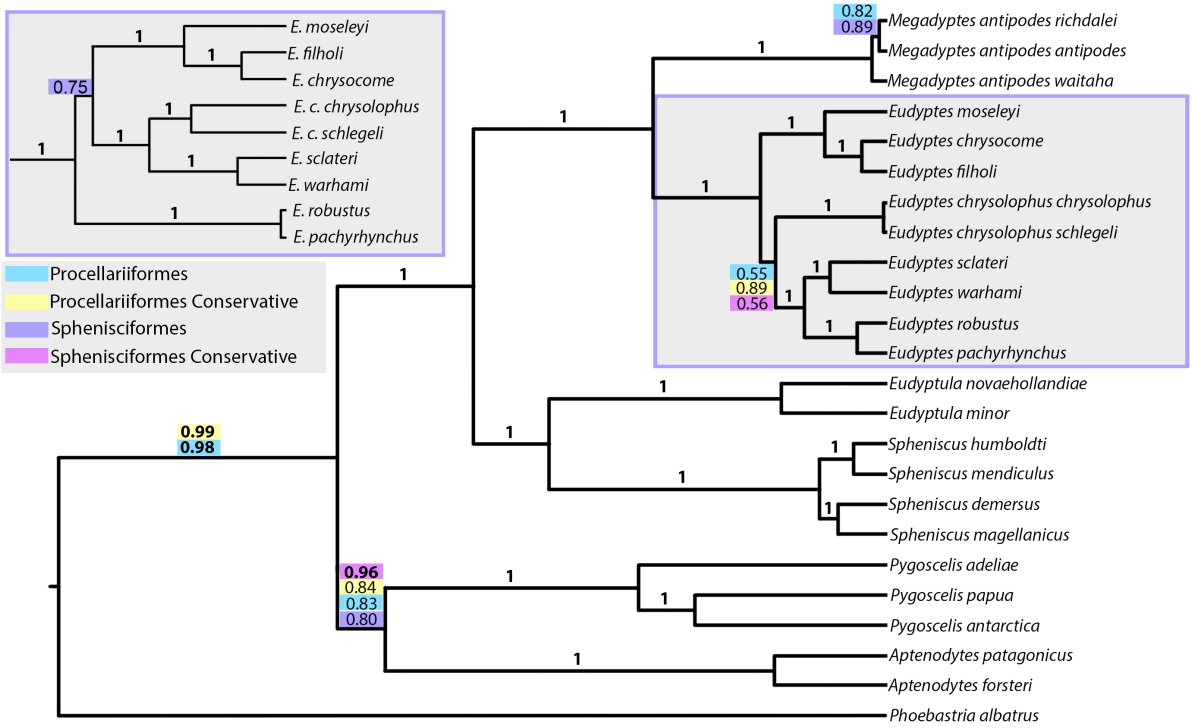
Supplementary Figure 5. MapDamage reports for the final round of mapping of six partial ancient mitogenomes to the *Eudyptes sclateri* mitogenome. The top panels show the characteristic high frequency of purines (A and G) immediately before the reads. The lower panels show the accumulation of 5' C to T (red) and 3' G to A (blue) misincorporations characteristic of aDNA.



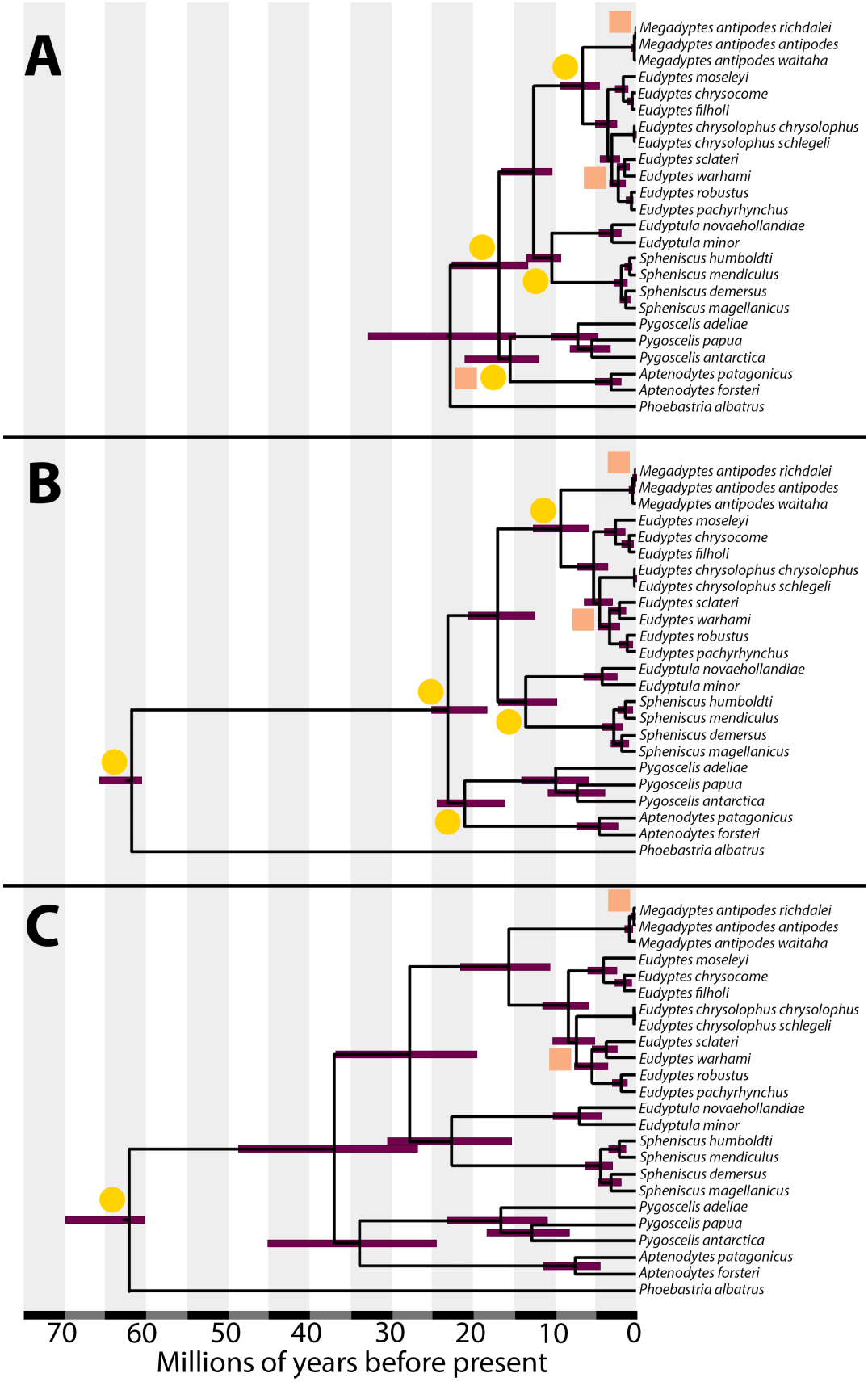
Supplementary Figure 6. MapDamage reports for the final round of mapping of five partial ancient mitogenomes to the *Eudyptes sclateri* mitogenome. The top panels show the characteristic high frequency of purines (A and G) immediately before the reads. The lower panels show the accumulation of 5' C to T (red) and 3' G to A (blue) misincorporations characteristic of aDNA.



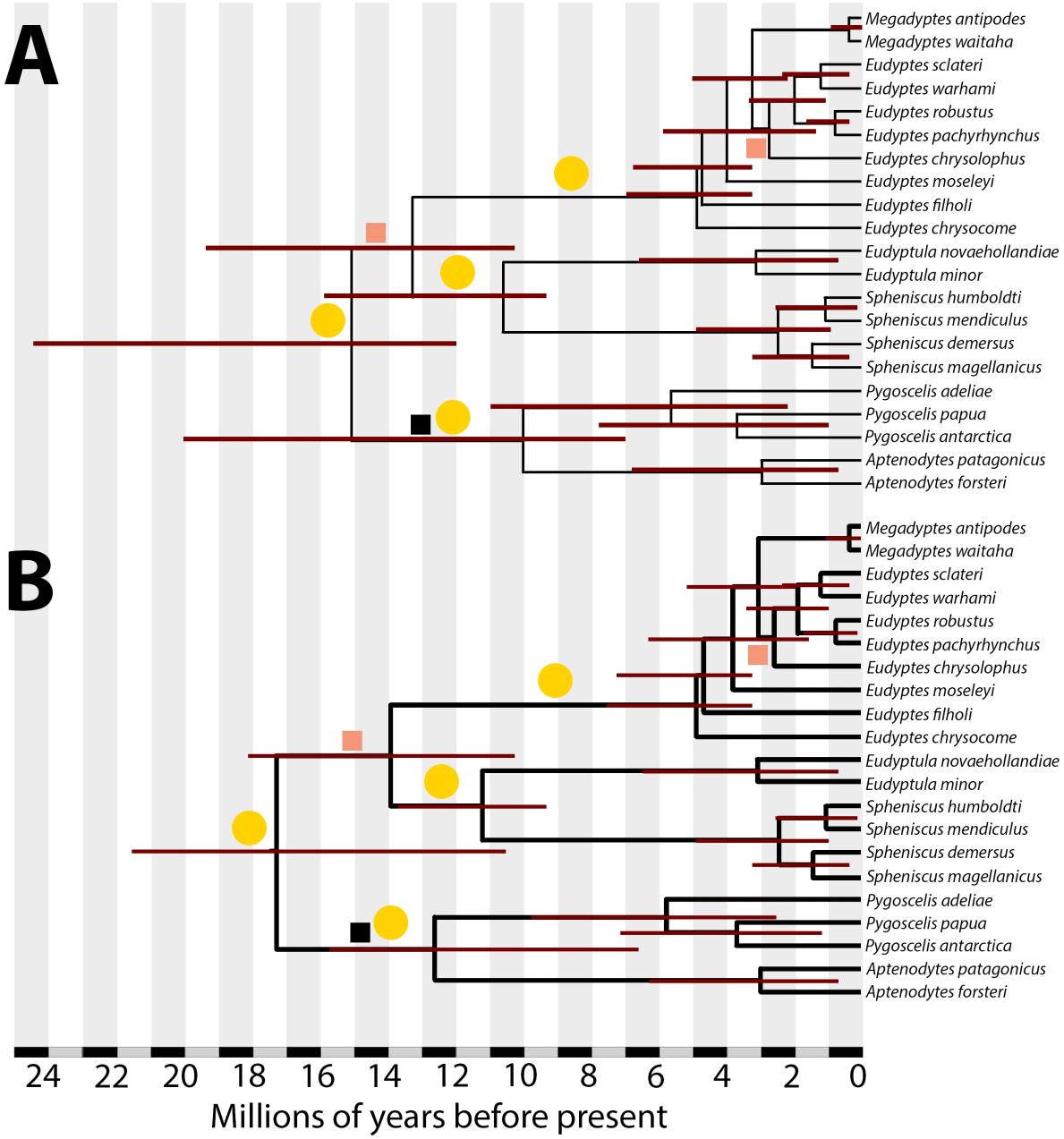
Supplementary Figure 7. Uncalibrated phylogenetic tree(s) across all modern penguin taxa inferred from 36 mitogenomes (Sphenisciformes only) and 37 mitogenomes (Sphenisciformes and Procellariiformes). The main phylogeny was run with *Phoebastria albatrus* as the outgroup, with posterior probabilities represented below each branch (blue). Posterior probabilities above each branch (purple) was achieved when the phylogeny was run without *P. albatrus*, and therefore there is no posterior probability for Sphenisciform/Procellariform divergence for that analysis. Large black posterior probabilities (also above the line), were obtained from both analyses. Inset A shows those posterior probabilities that could not be visually incorporated into the main figure, for the clade representing *Eudyptes warhami*. Inset B is taken directly from the B/C box, and can be compared with Inset C, which shows posterior probability and topology differences (specifically for AD419) when the analysis was conducted without *Phoebastria albatrus*.



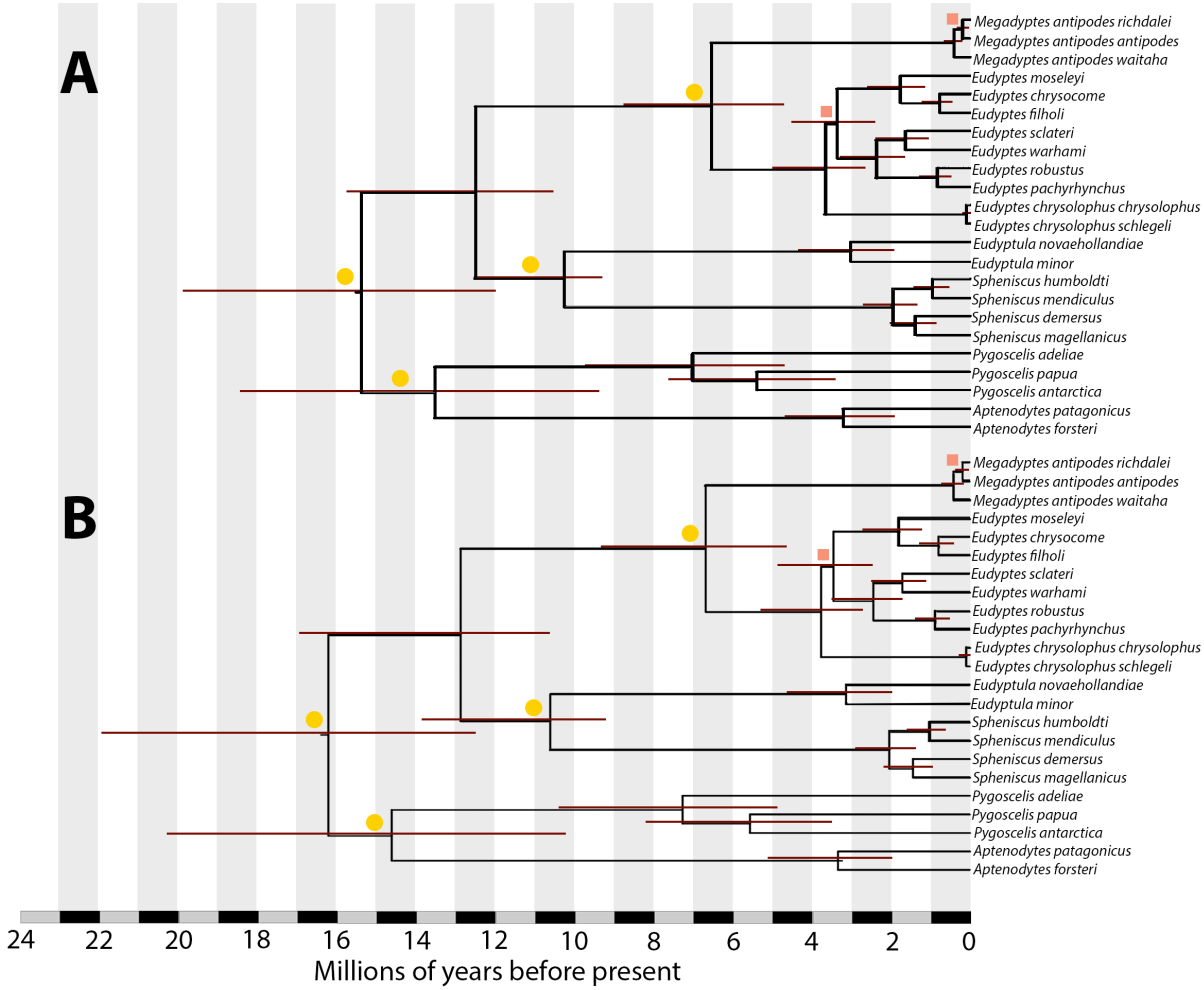
Supplementary Figure 8. Uncalibrated phylogenetic trees(s) across all modern penguin taxa inferred from up to 24 of the highest-quality mitogenomes, using four different alignments. 1) one representative of each taxa, with *Phoebastria albatrus* as the outgroup (blue); 2) one representative of each taxa, where we consider *Megadyptes antipodes antipodes*/*M. a. richdalei* as conspecific and *Eudyptes chrysolophus chrysolophus*/*E. c. schlegeli* as conspecific, with *Phoebastria albatrus* as outgroup (yellow); 3) one representative of each Sphenisciform taxon only (purple); and 4) one representative of each Sphenisciform taxon only, where we consider *M. antipodes antipodes*/*M. antipodes richdalei* as conspecific and *Eudyptes chrysolophus chrysolophus*/*E. c. schlegeli* as conspecific. Posterior probabilities of 1 were consistent for all analyses. The only inconsistency in the topology was the placement of *Eudyptes chrysolophus chrysolophus* and *E chrysolophus schlegeli* within *Eudyptes*, and this is shown in the inset.



Supplementary Figure 9. Calibrated phylogenetic trees across all modern penguin taxa inferred by the highest-quality mitogenomes using *Phoebastria albatrus* as an outgroup. The yellow circles show the nodes that were calibrated in each analysis. Posterior probabilities were all >0.99 except when indicated with the orange square. 95% HPD error bars are shown for each node. A) has all calibration points except *Waimanu manneringi*; B) has all calibration points including *W. manneringi*; and C) has only the *W. manneringi* calibration point.



Supplementary Figure 10. Calibrated phylogenetic trees across all modern penguin taxa using a conservative approach, where we consider *Eudyptes chrysolophus chrysolophus* and *E. c. schlegeli* and *Megadyptes antipodes antipodes* and *M. a. richdalei* as conspecific. **A)** was run with a birth-death speciation prior; and **B)** was run with calibrated Yule speciation prior. Posterior probabilities were >0.95 except when indicated with the orange square (*Spheniscus* + *Eudyptula* + *Eudyptes* + *Megadyptes* group, A: 0.46, B: 0.62; *Eudyptes chrysolophus schlegeli*, A: 0.93, B: 0.94). When we analysed the same samples as an uncalibrated phylogeny, *Aptenodytes*/*Pygoscelis* had a lower posterior probability of 0.85 (indicated by the black square), and *Eudyptes chrysolophus schlegeli* was recovered as basal to all *Eudyptes* (posterior probability: 0.95, not shown). The yellow circles show the nodes that were calibrated in each analysis. 95% HPD error bars are shown for each node.



Supplementary Figure 11. Calibrated phylogenetic trees across all modern penguin taxa inferred by the highest-quality mitogenomes for two evolutionary models. **A)** was run with a birth-death speciation prior; and **B)** was run with calibrated Yule speciation prior. Fossil calibrations are marked with a yellow circle. Posterior probabilities were all >0.99, except when indicated with the orange square (*Megadyptes antipodes antipodes*/*M. a. richdalei* group, A: 0.88, B: 0.88; *Eudyptes,* A: 0.75, B: 0.74). 95% HPD error bars are shown for each node.

Supplementary Figure 12. M*egadyptes antipodes richdalei* paratype. Unregistered Canterbury Museum specimen (sample ACAD12997). Photo Jamie Wood.

Supplementary Table 1. *Eudyptes*, *Megadyptes* and *Aptenodytes* samples used for obtaining COI, CR and near-complete mitogenome sequences. Details include the the sample type, the collection locality, the number of base pairs obtained for the four molecular markers and the taxon as indicated by molecular and morphological comparisons. 'sp.' indicates that we could only assign the sample to *Eudyptes*, *Megadyptes* or *Eudyptes*/*Megadyptes*. 'X' indicates that the marker of interest did not amplify, or the taxon is unknown. Those samples with a museum number beginning with 'Av' are from Canterbury Museum, those with 'OR' 'DM' or 'S' are from NMNZ, and those with AM are from the Auckland War Memorial Museum. 'GB' indicates that the particular sequence is already available on GenBank, yet additional sequences from a different molecular marker has been obtained in this study (see also Cole *et al*. 2019a; *Chapter 4* and Cole *et al*. 2018; *Chapter 5*). When two samples are recorded for the same individual, it indicates that multiple DNA extractions have been performed.

Sample Number	Museum Number	Sample Type	Locality	Genus	Taxon	COI (bp)	CR *Eudyptes* (bp)	CR *Megadyptes* (bp)	Mitogenome
AD82	Av26955	Bone	Chatham Islands	*Eudyptes* or *Megadyptes*	sp.	X	X	X	X
AD83	Av27126	Bone	Chatham Islands	*Eudyptes* or *Megadyptes*	sp.	X	X	X	X
AD84	Av7661	Bone	Chatham Islands	*Eudyptes* or *Megadyptes*	sp.	X	X	X	X
AD85	Av7655	Bone	Chatham Islands	*Eudyptes*	*warhami*	142	X	X	X
AD86	Av7654	Bone	Chatham Islands	*Eudyptes*	*sclateri*	356	131	X	X
AD87	Av5682	Bone	Chatham Islands	*Eudyptes* or *Megadyptes*	sp.	X	X	X	X
AD88	S.45876	Bone	Chatham Islands	*Megadyptes*	*antipodes richdalei*	133	X	X	2
AD91	S.35913	Bone	Native Island, Stewart Island	*Megadyptes*	*antipodes waitaha*	141 GB	X	X	1
AD93	Otago UF2/L2 LB3 A6 Bird, L2 + Llb Bag, LB/D/F	Bone	Long Beach, Otago, NZ	*Megadyptes*	*antipodes antipodes*	377 GB	X	250(GB)	1
AD94	Otago UE3/L2	Bone	Long Beach, Otago, NZ	*Megadyptes*	*antipodes antipodes*	141 GB	X	250(GB)	1
AD95	S.26921	Bone	Chatham Islands	*Megadyptes*	*antipodes richdalei*	141	X	250	X
AD156	S.25157	Bone	Chatham Islands	*Eudyptes*	*warhami*	283	X	X	1
AD157	Av6816	Bone	Chatham Islands	*Eudyptes*	*warhami*	341	X	X	1
AD158	S.27259	Bone	Chatham Islands	*Eudyptes*	*warhami*	378	X	X	2
AD159	S.47918	Bone	Chatham Islands	*Megadyptes*	sp.	234	X	X	X
AD161	S.33007	Bone	Chatham Islands	*Eudyptes*	*warhami*	141	X	X	2

Sample Number	Museum Number	Sample Type	Locality	Genus	Taxon	COI (bp)	CR *Eudyptes* (bp)	CR *Megadyptes* (bp)	Mitogenome
AD162	S.47917	Bone	Chatham Islands	*Eudyptes*	*warhami*	356	X	X	1
AD163	S.26908	Bone	Chatham Islands	*Eudyptes* or *Megadyptes*	sp.	X	X	X	X
AD175/AD420	OR.12822	Tissue	South Thule Island, South Sandwich Islands	*Eudyptes*	*chrysolophus chrysolophus*	498 GB	X	X	X
AD176/AD419	OR.13284	Tissue	Macquarie Island	*Eudyptes*	*chrysolophus schlegeli*	483 GB	X	X	1
AD177/AD417	OR.13285	Tissue	Macquarie Island	*Eudyptes*	*chrysolophus schlegeli*	498 GB	X	X	1
AD178/AD416	OR.13282	Tissue	Macquarie Island	*Eudyptes*	*chrysolophus schlegeli*	499 GB	X	X	1
AD179/AD421	OR.10287	Tissue	Cape Hallett, Antarctica	*Eudyptes*	*chrysolophus* ssp.	499 GB	X	X	X
AD180/AD418	OR.12821	Tissue	South Thule Island, South Sandwich Islands	*Eudyptes*	*chrysolophus chrysolophus*	423 GB	X	X	X
AD181/AD422	OR.10286	Tissue	Sabrina Island	*Eudyptes*	*chrysolophus* ssp.	499 GB	X	X	X
AD182/AD415	OR.13286	Tissue	Macquarie Island	*Eudyptes*	*chrysolophus schlegeli*	355 GB	X	X	1
AD207	S.47919	Bone	Chatham Islands	*Eudyptes*	*sclateri*	142	X	X	1
AD248	Av21751	Bone	Rakauteru Cave, Kaikoura, Canterbury, NZ	*Eudyptes*	*warhami*	286 GB	X	X	X
AD250	Av27867	Bone	Te One, Chatham Islands	*Eudyptes*	*warhami*	356	131	X	X
AD251	Av5681	Bone	Chatham Islands	*Eudyptes*	sp.	338	X	X	X
AD252	Av11287	Bone	Chatham Islands	*Megadyptes*	*antipodes richdalei*	X	X	250	1
AD253	Av21761	Bone	Campbell Island	*Eudyptes*	*filholi*	285 GB	X	X	1
AD266	OR.1067	Toe-pad	Solander Island, Southland, NZ	*Eudyptes*	*pachyrhynchus*	452 GB	X	X	1
AD270	OR.19353	Toe-pad	The Snares	*Eudyptes*	*robustus*	498 GB	X	X	1
AD289	S.42136	Bone	Delaware Bay, Nelson, NZ	*Megadyptes*	*antipodes waitaha*	109 GB	X	X	1
AD290	S.47920	Bone	Maunganui Beach, Chatham Islands	*Eudyptes*	*warhami*	348	X	X	X
AD291	S.24277	Bone	Chatham Islands	*Eudyptes*	*warhami*	323	X	X	X
AD293	S.36229.1	Bone	Mangere Island, Chatham Islands	*Eudyptes* or *Megadyptes*	sp.	X	X	X	X

Sample Number	Museum Number	Sample Type	Locality	Genus	Taxon	COI (bp)	CR *Eudyptes* (bp)	CR *Megadyptes* (bp)	Mitogenome
AD296	S.36837	Bone	Pitt Island, Chatham Islands	*Eudyptes* or *Megadyptes*	sp.	X	X	X	X
AD297	S.37744	Bone	Pitt Island, Chatham Islands	*Eudyptes*	*sclateri*	356	131	X	X
AD299	AM-12072	Bone	North West Pitt Island, Chatham Islands	*Eudyptes*	*warhami*	356	131	X	X
AD300	AM-12067	Bone	Waitangi, Chatham Islands	*Eudyptes* or *Megadyptes*	sp.	X	X	X	X
AD301	AM-12063	Bone	North West Pitt Island, Chatham Islands	*Megadyptes*	*antipodes richdalei*	141	X	250	X
AD302	OR.12646	Toe-pad	Campbell Island	*Eudyptes*	*sclateri*	499 (GB)	X	X	1
AD309	3/2/8 Black Rocks Midden BR3 N168-9/77	Bone	Black Rocks, Wairarapa, NZ	*Eudyptes*	*warhami*	356 (GB)	X	X	1
AD342	Av25695	Bone	Le Bons Bay, Banks Peninsula, Canterbury, NZ	*Eudyptes*	*warhami*	499 (GB)	X	X	1
AD345	Av27407	Bone	Chatham Islands	*Eudyptes*	*warhami*	356	X	X	X
AD346	Av27612	Bone	Chatham Islands	*Eudyptes* or *Megadyptes*	sp.	X	X	X	X
AD347	Av27227	Bone	Chatham Islands	*Eudyptes* or *Megadyptes*	sp.	X	X	X	X
AD348	Av26955	Bone	Chatham Islands	*Eudyptes* or *Megadyptes*	sp.	X	X	X	X
AD349	Av27823	Bone	Chatham Islands	*Eudyptes* or *Megadyptes*	sp.	X	X	X	X
AD352	S.47921	Bone	Chatham Islands	*Eudyptes* or *Megadyptes*	sp.	X	X	X	X
AD354	AM-14154	Bone	Chatham Islands	*Eudyptes*	*warhami*	356	X	X	X
AD370	Av27126	Bone	Long Beach, Chatham Islands	*Eudyptes* or *Megadyptes*	sp.	X	X	X	X
ACAD12997	Unregistered Canterbury Museum	Bone	Te One; Chatham Islands	*Megadyptes*	*antipodes richdalei*	X	X	X	2
JW820	MPPEI1A	Blood	RSA Point, Marion Island	*Eudyptes*	*chrysolophus chrysolophus*	X	X	X	1
JW868	4458A	Blood	Green Gorge, Macquarie Island	*Eudyptes*	*chrysolophus schlegeli*	X	X	X	1
NRP118-1	NRP118-1	Blood	Amsterdam Island	*Eudyptes*	*moseleyi*	X	X	X	1
KPFORT001	KPFORT001	Blood	Fortuna Bay, South Georgia	*Aptenodytes*	*patagonicus*	X	X	X	1

Supplementary Table 2. Sequences obtained from GenBank. Sample Number indicates that additional sequence data is available from that individual (see also *Chapter 3*, Cole *et al*. 2019a; *Chapter 4* and Cole *et al*. 2018; *Chapter 5* and *Supplementary Table 1*). 'X' indicates that no additional sample is available for that individual.

Order	Species	Accession Number	Sequence	Sample Number
Sphenisciformes	*Pygoscelis adeliae*	DQ137183	COI	X
Sphenisciformes	*Pygoscelis adeliae*	GQ925801	Mitogenome	X
Sphenisciformes	*Pygoscelis antarctica*	EU525474	COI	X
Sphenisciformes	*Pygoscelis antarctica*	KU356673	Mitogenome	X
Sphenisciformes	*Pygoscelis papua*	EU525486	COI	X
Sphenisciformes	*Pygoscelis papua*	KU356677	Mitogenome	X
Sphenisciformes	*Aptenodytes forsteri*	EU525299	COI	X
Sphenisciformes	*Aptenodytes forsteri*	KT159230	Mitogenome	X
Sphenisciformes	*Aptenodytes patagonicus*	EU525303	COI	X
Sphenisciformes	*Spheniscus demersus*	EU525504	COI	X
Sphenisciformes	*Spheniscus demersus*	KC914350	Mitogenome	X
Sphenisciformes	*Spheniscus humboldti*	DQ137180	COI	X
Sphenisciformes	*Spheniscus humboldti*	KM891593	Mitogenome	X
Sphenisciformes	*Spheniscus magellanicus*	EU525505	COI	X
Sphenisciformes	*Spheniscus magellanicus*	KU361803	Mitogenome	X
Sphenisciformes	*Spheniscus mendiculus*	DQ137179	COI	X
Sphenisciformes	*Spheniscus mendiculus*	KU361807	Mitogenome	X
Sphenisciformes	*Eudyptula minor*	EU525402	COI	X
Sphenisciformes	*Eudyptula minor*	AF362763	Mitogenome	X
Sphenisciformes	*Eudyptula novaehollandiae*	EU525396	COI	X
Sphenisciformes	*Eudyptula novaehollandiae*	MF370525	Mitogenome	X
Sphenisciformes	*Megadyptes antipodes antipodes*	DQ137184	COI	X
Sphenisciformes	*Megadyptes antipodes antipodes*	MK120819	COI	AD93
Sphenisciformes	*Megadyptes antipodes antipodes*	KP644149	CR	AD93
Sphenisciformes	*Megadyptes antipodes antipodes*	MK120821	COI	AD94

Order	Species	Accession Number	Sequence	Sample Number
Sphenisciformes	*Megadyptes antipodes antipodes*	FJ391951	CR	AD94
Sphenisciformes	*Megadyptes antipodes antipodes*	MK120820	CR	AD96
Sphenisciformes	*Megadyptes antipodes antipodes*	FJ391943	CR	X
Sphenisciformes	*Megadyptes antipodes antipodes*	FJ391944	CR	X
Sphenisciformes	*Megadyptes antipodes antipodes*	FJ391945	CR	X
Sphenisciformes	*Megadyptes antipodes antipodes*	FJ391946	CR	X
Sphenisciformes	*Megadyptes antipodes antipodes*	FJ391947	CR	X
Sphenisciformes	*Megadyptes antipodes antipodes*	FJ391948	CR	X
Sphenisciformes	*Megadyptes antipodes antipodes*	FJ391949	CR	X
Sphenisciformes	*Megadyptes antipodes antipodes*	FJ391950	CR	X
Sphenisciformes	*Megadyptes antipodes antipodes*	FJ391951	CR	X
Sphenisciformes	*Megadyptes antipodes antipodes*	FJ391952	CR	X
Sphenisciformes	*Megadyptes antipodes antipodes*	FJ391953	CR	X
Sphenisciformes	*Megadyptes antipodes antipodes*	FJ822137	CR	X
Sphenisciformes	*Megadyptes antipodes antipodes*	FJ822138	CR	X
Sphenisciformes	*Megadyptes antipodes antipodes*	FJ822139	CR	X
Sphenisciformes	*Megadyptes antipodes antipodes*	FJ822141	CR	X
Sphenisciformes	*Megadyptes antipodes waitaha*	MK120822	COI	AD91
Sphenisciformes	*Megadyptes antipodes waitaha*	FJ391965	CR	AD91
Sphenisciformes	*Megadyptes antipodes waitaha*	MK120828	COI	AD289
Sphenisciformes	*Megadyptes antipodes waitaha*	FJ391964	CR	AD92
Sphenisciformes	*Megadyptes antipodes waitaha*	MK120773	CR	AD276
Sphenisciformes	*Megadyptes antipodes waitaha*	MK120768	CR	AD283
Sphenisciformes	*Megadyptes antipodes waitaha*	MK120765	CR	AD284
Sphenisciformes	*Megadyptes antipodes waitaha*	MK120770	CR	AD285
Sphenisciformes	*Megadyptes antipodes waitaha*	MK120766	CR	AD286
Sphenisciformes	*Megadyptes antipodes waitaha*	MK120771	CR	AD371
Sphenisciformes	*Megadyptes antipodes waitaha*	MK120772	CR	AD375

Order	Species	Accession Number	Sequence	Sample Number
Sphenisciformes	*Megadyptes antipodes waitaha*	MK120767	CR	AD378
Sphenisciformes	*Megadyptes antipodes waitaha*	KP644153	CR	X
Sphenisciformes	*Megadyptes antipodes waitaha*	FJ391950	CR	X
Sphenisciformes	*Megadyptes antipodes waitaha*	FJ391954	CR	X
Sphenisciformes	*Megadyptes antipodes waitaha*	FJ391955	CR	X
Sphenisciformes	*Megadyptes antipodes waitaha*	FJ391956	CR	X
Sphenisciformes	*Megadyptes antipodes waitaha*	FJ391957	CR	X
Sphenisciformes	*Megadyptes antipodes waitaha*	FJ391958	CR	X
Sphenisciformes	*Megadyptes antipodes waitaha*	FJ391959	CR	X
Sphenisciformes	*Megadyptes antipodes waitaha*	FJ391960	CR	X
Sphenisciformes	*Megadyptes antipodes waitaha*	FJ391961	CR	X
Sphenisciformes	*Megadyptes antipodes waitaha*	FJ391962	CR	X
Sphenisciformes	*Megadyptes antipodes waitaha*	FJ391963	CR	X
Sphenisciformes	*Megadyptes antipodes waitaha*	FJ391964	CR	X
Sphenisciformes	*Megadyptes antipodes waitaha*	FJ391965	CR	X
Sphenisciformes	*Megadyptes antipodes waitaha*	FJ391966	CR	X
Sphenisciformes	*Eudyptes chrysocome*	DQ525800	COI	X
Sphenisciformes	*Eudyptes chrysocome*	AP009189	Mitogenome	X
Sphenisciformes	*Eudyptes chrysolophus chrysolophus*	MF469857	COI	AD175
Sphenisciformes	*Eudyptes chrysolophus chrysolophus*	MF469863	COI	AD179
Sphenisciformes	*Eudyptes chrysolophus chrysolophus*	MF469859	COI	AD180
Sphenisciformes	*Eudyptes chrysolophus chrysolophus*	MF469864	COI	AD181
Sphenisciformes	*Eudyptes chrysolophus chrysolophus*	FJ582595	COI	X
Sphenisciformes	*Eudyptes filholi*	MF469867	COI	AD253
Sphenisciformes	*Eudyptes filholi*	DQ525795	COI	X
Sphenisciformes	*Eudyptes moseleyi*	DQ525790	COI	X
Sphenisciformes	*Eudyptes pachyrhynchus*	MF469877	COI	AD266
Sphenisciformes	*Eudyptes pachyrhynchus*	EU525345	COI	X
Sphenisciformes	*Eudyptes pachyrhynchus*	FJ391950	CR	X

Order	Species	Accession Number	Sequence	Sample Number
Sphenisciformes	*Eudyptes robustus*	MF469856	COI	AD270
Sphenisciformes	*Eudyptes robustus*	EU525347	COI	X
Sphenisciformes	*Eudyptes chrysolophus schlegeli*	MF469865	COI	AD176
Sphenisciformes	*Eudyptes chrysolophus schlegeli*	MF469862	COI	AD177
Sphenisciformes	*Eudyptes chrysolophus schlegeli*	MF469861	COI	AD178
Sphenisciformes	*Eudyptes chrysolophus schlegeli*	MF469858	COI	AD182
Sphenisciformes	*Eudyptes chrysolophus schlegeli*	FJ582599	COI	X
Sphenisciformes	*Eudyptes sclateri*	MF469866	COI	AD304
Sphenisciformes	*Eudyptes sclateri*	MK120811	COI	AD302
Sphenisciformes	*Eudyptes warhami*	MK120815	COI	AD248
Sphenisciformes	*Eudyptes warhami*	MK120818	COI	AD282
Sphenisciformes	*Eudyptes warhami*	MK120814	COI	AD309
Sphenisciformes	*Eudyptes warhami*	MK120813	COI	AD342
Sphenisciformes	*Eudyptes warhami*	MK120812	COI	AD391
Sphenisciformes	*Eudyptes warhami*	MK120817	COI	AD394
Sphenisciformes	*Eudyptes warhami*	MK120809	COI	AD389
Procellariiformes	*Phoebastria albatrus*	KJ735514	Mitogenome	X
Procellariiformes	*Diomedia exulans*	DQ137168	COI	X
Procellariiformes	*Procellaria cinerea*	AP009191	Mitogenome	X
Procellariiformes	*Pterodroma brevirostris*	AY158678	Mitogenome	X
Gaviiformes	*Gavia stellata*	AY293618	Mitogenome	X

Supplementary Table 3. Details associated with the quality and quantity of the ancient mitogenomes. All enriched mitogenomes were prepared at ACAD.

Sample Number	Museum Number	Taxon	Total Filtered Reads	Unique Aligned Reads	Sequence Length (bp)	% of Coverage	Mean Sequencing Depth
AD88a	S.45876	*Megadyptes antipodes richdalei*	59785	474	15592	50.6	0.98
AD88b	S.45876	*Megadyptes antipodes richdalei*	197782	256	15597	84.4	2.14
AD91	S.35913	*Megadyptes antipodes waitaha*	51461	10714	15568	100	46.99
AD93	Otago UF2/L2 LB3 A6 Bird, L2+ Llb Bag, LB/D/F	*Megadyptes antipodes antipodes*	523134	251733	15568	100	1401.4
AD94	Otago UE3/L2	*Megadyptes antipodes antipodes*	47824	2870	15570	100	14.46
AD95	S.26921	*Megadyptes antipodes richdalei*	24844	14213	15566	100	72.96
AD156	S.25157	*Eudyptes warhami*	19575	146	15574	45	0.57
AD157	Av6816	*Eudyptes warhami*	299152	225	15566	49.2	1.02
AD161a	S.33007	*Eudyptes warhami*	4773	115	15574	24.7	0.32
AD161b	S.33007	*Eudyptes warhami*	23862	76	15566	34.5	0.45
AD162	S.47917	*Eudyptes warhami*	26281	102	15574	42.9	0.52
AD253	Av21761	*Eudyptes filholi*	51319	44744	15559	100	221.97
AD266	OR.1067	*Eudyptes pachyrhynchus*	65458	54999	15565	99.6	251.41
AD270	OR.19353	*Eudyptes robustus*	31593	29402	15564	99.9	185.6
AD289	S.42136	*Megadyptes antipodes waitaha*	44149	6085	15550	100	28.27
AD302	OR.12646	*Eudyptes sclateri*	53514	51676	15555	100	271.63
AD309	S/2/8 Black Rocks Midden BR3	*Eudyptes warhami*	85558	6973	15563	100	32.97
AD342	Av25695	*Eudyptes warhami*	43857	30836	15562	100	222.99
AD415	OR.13286	*Eudyptes chrysolophus schlegeli*	41086	37180	15560	99.3	196.17
AD416	OR.13282	*Eudyptes chrysolophus schlegeli*	687771	407245	15560	99.6	2226.01
AD417	OR.13285	*Eudyptes chrysolophus schlegeli*	46822	43052	15562	99.3	232.44
AD419	OR.13284	*Eudyptes chrysolophus schlegeli*	654399	421344	15560	99.6	2120.04
ACAD12997	Unregistered Canterbury Museum	*Megadyptes antipodes richdalei*	49866	44185	15566	100	230.05

Supplementary Table 4. Partitioning schemes and substitution models for nine different alignments of the mitogenomes, determined using Partition Finder2 (Lanfear *et al*., 2016). Alignment indicates which samples were used; 'Best' included the highest quality genome per taxon, for *Eudyptes chrysolophus schlegeli* we use AD419, for *E. warhami* we use AD309, for *Megadyptes antipodes antipodes* we use AD94, for *M. a. richdalei* we use ACAD12997 and for *M. a. waitaha* we use AD289; 'Conservative' indicates that *Eudyptes chrysolophus chrysolophus* and *E. c. schlegeli* are considered conspecific and *Megadyptes antipodes antipodes* and *M. a. richdalei* are considered conspecific; 'All' included all mitogenomes.

Alignment	Outgroup	Partition Number	Composition	Substitution Model
Best	Procellariiformes	1	COI_1, CO2_1, CO3_1	TRNEF+I
Best	Procellariiformes	2	COI_2, CO2_2, CO3_2	HKY+I
Best	Procellariiformes	3	Cytb_1, ND1_1, ND3_2, ND4L_1, tRNAs, 12SrRNA	TVM+I+G
Best	Procellariiformes	4	ATP6_2, ATP8_2, Cytb_2, ND1_2, ND2_2, ND4_2, ND4L_2, ND5_2	TVM+I+G
Best	Procellariiformes	5	ATP8_1, ND2_1, ND4_1, ND5_1, 16SrRNA	K81UF+I+G
Best	Procellariiformes	6	ATP6_1, ND3_1, ND6_1, ND6_2	HKY+I+G
Best	Procellariiformes	7	Non-coding	HKY+G
Best	Procellariiformes and Gaviiformes	1	COI_1, CO2_1, CO3_1	TRNEF+I
Best	Procellariiformes and Gaviiformes	2	Cytb_2, COI_2, CO2_2, CO3_2, ND1_2, ND4L_2	HKY+I
Best	Procellariiformes and Gaviiformes	3	ATP8_1, Cytb1, ND1_1, ND2_1, ND4_1, ND4L_1, ND5_1, tRNAs, 12SrRNA, 16SrRNA	GTR+I+G
Best	Procellariiformes and Gaviiformes	4	ATP6_2, ATP8_2, ND2_2, ND3_2, ND4_2, ND5_2	TVM+I+G
Best	Procellariiformes and Gaviiformes	5	ATP6_1, ND3_1, ND6_1, ND6_2	HKY+I+G
Best	Procellariiformes and Gaviiformes	6	Non-coding	HKY+G
Best	*Aptenodytes* and *Pygoscelis*	1	COI_1, CO2_1, CO3_1	TRNEF+I
Best	*Aptenodytes* and *Pygoscelis*	2	ATP8_2, Cytb_2, CO1_2, CO2_2, CO3_2, ND1_2, ND4L_2	HKY+I
Best	*Aptenodytes* and *Pygoscelis*	3	Cytb_1, ND1_1, ND3_2, ND4L_2, tRNAs, 12SrRNA, 16SrRNA	TVM+I+G
Best	*Aptenodytes* and *Pygoscelis*	4	ATP6_1, ATP8_1, ND2_1, ND3_1, ND4_1, ND5_1, ND6_2	HKY+I+G
Best	*Aptenodytes* and *Pygoscelis*	5	ATP6_2, ND2_2, ND4_2, ND5_2	HKY+I+G
Best	*Aptenodytes* and *Pygoscelis*	6	ND6_1	TVM
Best	*Aptenodytes* and *Pygoscelis*	7	Non-coding	HKY+G
Conservative	Procellariiformes	1	COI_1, CO2_1, CO3_1	TRNEF+I
Conservative	Procellariiformes	2	CO1_2, CO2_2, CO3_2	HKY+I
Conservative	Procellariiformes	3	Cytb_1, ND1_1, ND3_2, ND4L_1, tRNAs, 12SrRNA	TVM+I+G
Conservative	Procellariiformes	4	ATP6_2, ATP8_2, Cytb_2, ND1_2, ND2_2, ND4_2, ND4L_2, ND5_2	TVM+I+G

Alignment	Outgroup	Partition Number	Composition	Substitution Model
Conservative	Procellariiformes	5	ATP8_1, ND2_1, ND4_1, ND5_1, 16SrRNA	K81UF+I+G
Conservative	Procellariiformes	6	ATP6_1, ND3_1, ND6_1, ND6_2	HKY+I+G
Conservative	Procellariiformes	7	Non-coding	HKY+G
Conservative	Procellariiformes and Gaviiformes	1	COI_1, CO2_1, CO3_1	TRNEF+I
Conservative	Procellariiformes and Gaviiformes	2	CO1_2, CO2_2, CO3_2	HKY+I
Conservative	Procellariiformes and Gaviiformes	3	Cytb_1, ND1_1, ND3_2, ND4L_1, tRNAs, 12SrRNA	TVM+I+G
Conservative	Procellariiformes and Gaviiformes	4	ATP6_2, ATP8_2, Cytb_2, ND1_2, ND2_2, ND4_2, ND4L_2, ND5_2	TVM+I+G
Conservative	Procellariiformes and Gaviiformes	5	ATP8_1, ND2_1, ND4_1, ND5_1, 16SrRNS	K81UF+I+G
Conservative	Procellariiformes and Gaviiformes	6	ATP6_1, ND3_1, ND6_1, ND6_2	HKY+I+G
Conservative	Procellariiformes and Gaviiformes	7	Non-coding	HKY+G
Conservative	*Aptenodytes* and *Pygoscelis*	1	COI_1, CO2_1, CO3_1	TRNEF+I
Conservative	*Aptenodytes* and *Pygoscelis*	2	Cytb_2, CO1_2, CO2_2, CO3_2, ND1_2	HKY+I
Conservative	*Aptenodytes* and *Pygoscelis*	3	Cytb_1, ND3_2, ND4L_1, tRNAs, 12SrRNA, 16SrRNA	TVM+I+G
Conservative	*Aptenodytes* and *Pygoscelis*	4	ATP6_2, ATP8_1, ND2_1, ND3_1, ND4_1, ND5_1, ND6_1	HKY+I+G
Conservative	*Aptenodytes* and *Pygoscelis*	5	ATP6_2, ATP8_2, ND2_2, ND4_2, ND4L_2, ND5_2	HKY+I+G
Conservative	*Aptenodytes* and *Pygoscelis*	6	ND6_1	TVM+G
Conservative	*Aptenodytes* and *Pygoscelis*	7	Non-coding	HKY+G
All	Procellariiformes	1	COI_1, CO2_1, CO3_1	TRNEF+I
All	Procellariiformes	2	COI_2, CO2_2, CO3_2	HKY+I
All	Procellariiformes	3	Cytb_1, ND1_1, ND3_2, ND4L_1, tRNAs, 12SrRNA	TVM+I+G
All	Procellariiformes	4	ATP6_2, ATP8_2, Cytb_2, ND1_2, ND2_2, ND4_2, ND4L_2, ND5_2	TVM+I+G
All	Procellariiformes	5	ATP8_1, ND2_1, ND4_1, ND5_1, 16SrRNA	TIM+I+G
All	Procellariiformes	6	ATP6_1, ND3_1, ND6_1, ND6_2	HKY+I+G
All	Procellariiformes	7	Non-coding	HKY+G
All	Procellariiformes and Gaviiformes	1	COI_1, CO2_1, CO3_1	TRNEF+I
All	Procellariiformes and Gaviiformes	2	Cytb_2, COI_2, CO2_2, CO3_2	HKY+I
All	Procellariiformes and Gaviiformes	3	ATP8_1, Cytb_1, ND1_1, ND4_1, ND4L_1	K81UF+I+G
All	Procellariiformes and Gaviiformes	4	ATP8_2, ND1_2, ND2_2, ND3_2, ND4_2, ND4L_2, ND5_2	TVM+I+G

Alignment	Outgroup	Partition Number	Composition	Substitution Model
All	Procellariiformes and Gaviiformes	5	ATP6_1, ND3_1, ND6_1, ND6_2	HKY+I+G
All	Procellariiformes and Gaviiformes	6	ATP6_2, ND2_1, ND5_1, tRNAs, 12SrRNA, 16SrRNA	GTR+I+G
All	Procellariiformes and Gaviiformes	7	Non-coding	HKY+G
All	*Aptenodytes* and *Pygoscelis*	1	COI_1, CO2_1, CO3_1	TRNEF+I
All	*Aptenodytes* and *Pygoscelis*	2	ATP8_2, Cytb_2, CO1_2, CO2_2, CO3_2, ND1_2, ND4L_2	HKY+I
All	*Aptenodytes* and *Pygoscelis*	3	Cytb_1, ND1_1, ND3_2, ND4L_1, tRNAs	TVM+I+G
All	*Aptenodytes* and *Pygoscelis*	4	ATP6_1, ATP8_1, ND2_1, ND3_1, ND4_1, ND5_1, ND6_2, 12SrRNA, 16SrRNA	HKY+I+G
All	*Aptenodytes* and *Pygoscelis*	5	ATP6_2, ND2_2, ND4_2, ND5_2	HKY+I+G
All	*Aptenodytes* and *Pygoscelis*	6	ND6_1	TVM
All	*Aptenodytes* and *Pygoscelis*	7	Non-coding	HKY+G

Supplementary Table 5. Pairwise genetic distances obtained from 15755 bp of penguin mitogenomes. *Megadyptes antipodes antipodes* is AD94, *M. a. richdalei* is ACAD12997, *M. a. waitaha* is AD91, *Eudyptes chrysolophus schlegeli* is AD419 and *E. warhami* is AD309. The lower matrix is the % of pairwise genetic similarity, and the upper matrix is the number of nucleotide substitutions.

	Pygoscelis adeliae	*Pygoscelis papua*	*Pygoscelis antarctica*	*Aptenodytes forsteri*	*Aptenodytes patagonicus*	*Eudyptula novaehollandiae*	*Eudyptula minor*	*Spheniscus humboldti*	*Spheniscus mendiculus*	*Spheniscus demersus*	*Spheniscus magellanicus*	*Megadyptes antipodes antipodes*	*Megadyptes antipodes richdalei*	*Megadyptes antipodes waitaha*	*Eudyptes moseleyi*	*Eudyptes chrysocome*	*Eudyptes filholi*	*Eudyptes chrysolophus schlegeli*	*Eudyptes chrysolophus chrysolophus*	*Eudyptes sclateri*	*Eudyptes pachyrhynchus*	*Eudyptes robustus*	*Eudyptes warhami*
Pygoscelis adeliae	-	1098	1025	1658	2825	1695	1749	1674	1653	1630	1634	1663	1746	1786	1614	1624	1614	1664	1969	1618	1690	1660	1731
Pygoscelis papua	93	-	864	1656	2828	1720	1754	1686	1664	1665	1665	1642	1709	1745	1606	1601	1609	1695	1996	1615	1714	1669	1716
Pygoscelis antarctica	93.4	94.5	-	1634	2807	1709	1750	1668	1649	1640	1627	1634	1720	1754	1591	1593	1590	1660	1961	1615	1683	1634	1712
Aptenodytes forsteri	89.4	89.4	89.5	-	1871	1679	1718	1666	1645	1649	1629	1692	1758	1793	1628	1635	1633	1718	2023	1614	1700	1655	1697
Aptenodytes patagonicus	81.9	81.9	82	88	-	2802	2852	2799	2784	2777	2771	2835	2919	2902	2768	2779	2781	2780	2360	2757	2787	2771	2844
Eudyptula novaehollandiae	89.1	89	89.1	89.2	82.1	-	512	1304	1280	1288	1287	1437	1518	1546	1412	1406	1410	1464	1775	1416	1493	1456	1517
Eudyptula minor	88.8	88.8	88.8	89	81.7	96.7	-	1359	1336	1339	1344	1477	1557	1587	1427	1427	1433	1489	1797	1437	1499	1459	1521
Spheniscus humboldti	89.3	89.2	89.3	89.3	82.1	91.7	91.3	-	137	308	277	1392	1469	1505	1340	1361	1351	1413	1728	1346	1413	1372	1440
Spheniscus mendiculus	89.4	89.4	89.4	89.5	82.2	91.8	91.4	99.1	-	319	272	1368	1450	1485	1317	1332	1321	1391	1706	1331	1386	1345	1426
Spheniscus demersus	89.6	89.4	89.5	89.5	82.2	91.8	91.4	98	98	-	222	1364	1443	1475	1313	1332	1325	1378	1693	1314	1380	1344	1416
Spheniscus magellanicus	89.5	89.3	89.6	89.6	82.3	91.8	91.4	98.2	98.3	98.6	-	1364	1442	1474	1317	1335	1324	1378	1688	1317	1387	1348	1417
Megadyptes antipodes antipodes	89.4	89.5	89.6	89.2	81.8	90.8	90.6	91.1	91.3	91.3	91.3	-	172	237	943	941	926	1026	1323	912	1002	962	1034
Megadyptes antipodes richdalei	88.8	89.1	89	88.8	81.3	90.3	90.1	90.6	90.7	90.8	90.8	98.9	-	103	1033	1033	1018	1123	1426	1008	1094	1057	902
Megadyptes antipodes waitaha	88.7	88.9	88.9	88.6	81.4	90.2	89.9	90.5	90.6	90.7	90.7	98.5	99.4	-	1079	1077	1064	1166	1439	1049	1141	1100	945
Eudyptes moseleyi	89.6	89.7	89.8	89.6	82.2	90.9	90.8	91.4	91.6	91.6	91.6	94	93.4	93.2	-	314	302	591	924	533	587	551	658
Eudyptes chrysocome	89.6	89.7	89.8	89.5	82.2	91	90.8	91.3	91.5	91.5	91.4	94	93.4	93.2	98	-	128	588	916	522	583	556	652
Eudyptes filholi	89.7	89.7	89.8	89.5	82.2	91	90.8	91.3	91.5	91.5	91.5	94.1	93.5	93.3	98.1	99.2	-	576	910	508	567	539	639
Eudyptes chrysolophus schlegeli	89.5	89.3	89.5	89.1	82.2	90.7	90.6	91.1	91.2	91.3	91.3	93.6	92.9	92.7	96.3	96.3	96.4	-	507	557	570	565	711
Eudyptes chrysolophus chrysolophus	87.4	87.2	87.4	87	84.4	88.6	88.5	88.9	89.1	89.2	89.2	91.5	90.9	90.8	94.1	94.1	94.2	96.8	-	893	952	912	1039
Eudyptes sclateri	89.6	89.6	89.6	89.6	82.3	90.9	90.8	91.4	91.5	91.6	91.6	94.2	93.5	93.4	96.6	96.6	96.7	96.5	94.3	-	460	438	444
Eudyptes pachyrhynchus	89.3	89.1	89.3	89.2	82.1	90.6	90.5	91.1	91.2	91.3	91.2	93.7	93.1	92.9	96.4	96.4	96.5	96.4	94	97.2	-	199	601
Eudyptes robustus	89.4	89.4	89.6	89.5	82.2	90.7	90.7	91.3	91.5	91.5	91.4	93.9	93.3	93.1	96.5	96.5	96.6	96.5	94.2	97.3	98.7	-	578
Eudyptes warhami	88.9	89	89.1	89.1	81.8	90.3	90.3	90.8	90.9	91	90.9	93.4	94.2	94	95.8	95.8	95.9	95.6	93.3	97.2	96.3	96.4	-

Supplementary Table 6. Pairwise genetic distances obtained from 14117 bp of penguin mitogenomes, excluding all missing data. *Megadyptes antipodes antipodes* is AD94, *M. a. richdalei* is ACAD12997, *M. a. waitaha* is AD91, *Eudyptes chrysolophus schlegeli* is AD419 and *E. warhami* is AD309. The lower matrix is the % of pairwise genetic similarity, and the upper matrix is the number of nucleotide substitutions.

	Pygoscelis adeliae	*Pygoscelis papua*	*Pygoscelis antarctica*	*Aptenodytes forsteri*	*Aptenodytes patagonicus*	*Eudyptula novaehollandiae*	*Eudyptula minor*	*Spheniscus humboldti*	*Spheniscus mendiculus*	*Spheniscus demersus*	*Spheniscus magellanicus*	*Megadyptes antipodes antipodes*	*Megadyptes antipodes richdalei*	*Megadyptes antipodes waitaha*	*Eudyptes moseleyi*	*Eudyptes chrysocome*	*Eudyptes filholi*	*Eudyptes chrysolophus schlegeli*	*Eudyptes chrysolophus chrysolophus*	*Eudyptes sclateri*	*Eudyptes pachyrhynchus*	*Eudyptes robustus*	*Eudyptes warhami*
Pygoscelis adeliae	-	937	892	1396	1406	1391	1443	1375	1362	1352	1349	1384	1389	1387	1364	1377	1365	1354	1356	1369	1374	1369	1375
Pygoscelis papua	93.4	-	732	1382	1402	1409	1451	1387	1373	1377	1373	1359	1363	1359	1358	1356	1358	1380	1382	1357	1388	1372	1369
Pygoscelis antarctica	93.7	94.8	-	1372	1383	1400	1447	1368	1359	1358	1341	1373	1381	1373	1344	1349	1339	1348	1351	1362	1364	1350	1361
Aptenodytes forsteri	90.1	90.2	90.3	-	500	1370	1404	1364	1351	1353	1338	1416	1417	1417	1383	1385	1384	1395	1399	1363	1381	1362	1353
Aptenodytes patagonicus	90	90.1	90.2	96.5	-	1369	1417	1358	1350	1339	1334	1414	1417	1409	1360	1364	1370	1375	1379	1356	1378	1360	1353
Eudyptula novaehollandiae	90.1	90	90.1	90.3	90.3	-	415	1038	1022	1035	1029	1159	1165	1157	1161	1149	1150	1143	1149	1161	1176	1163	1160
Eudyptula minor	89.8	89.7	89.7	90.1	90	97.1	-	1087	1069	1079	1075	1200	1206	1199	1180	1176	1177	1172	1176	1185	1172	1162	1173
Spheniscus humboldti	90.3	90.2	90.3	90.3	90.4	92.6	92.3	-	119	237	237	1139	1145	1139	1092	1115	1106	1112	1120	1110	1102	1086	1110
Spheniscus mendiculus	90.4	90.3	90.4	90.4	90.4	92.8	92.4	99.2	-	247	239	1119	1125	1119	1076	1093	1083	1097	1105	1102	1082	1066	1093
Spheniscus demersus	90.4	90.2	90.4	90.4	90.5	92.7	92.4	98.3	98.3	-	163	1113	1119	1111	1083	1103	1095	1092	1099	1095	1086	1070	1092
Spheniscus magellanicus	90.4	90.3	90.5	90.5	90.6	92.7	92.4	98.3	98.3	98.8	-	1117	1123	1115	1076	1095	1085	1083	1090	1090	1084	1071	1087
Megadyptes antipodes antipodes	90.2	90.4	90.3	90	90	91.8	91.5	91.9	92.1	92.1	92.1	-	15	46	767	767	750	778	773	737	759	745	745
Megadyptes antipodes richdalei	90.2	90.3	90.2	90	90	91.7	91.5	91.9	92	92.1	92	99.9	-	39	770	770	753	782	777	742	762	750	745
Megadyptes antipodes waitaha	90.2	90.4	90.3	90	90	91.8	91.5	91.9	92.1	92.1	92.1	99.7	99.7	-	771	773	758	787	782	745	768	756	749
Eudyptes moseleyi	90.3	90.4	90.5	90.2	90.4	91.8	91.6	92.3	92.4	92.3	92.4	94.6	94.5	94.5	-	261	250	439	439	447	427	422	443
Eudyptes chrysocome	90.2	90.4	90.4	90.2	90.3	91.9	91.7	92.1	92.3	92.2	92.2	94.6	94.5	94.5	98.2	-	109	433	428	436	418	417	431
Eudyptes filholi	90.3	90.4	90.5	90.2	90.3	91.9	91.7	92.2	92.3	92.2	92.3	94.7	94.7	94.6	98.2	99.2	-	422	423	425	404	401	421
Eudyptes chrysolophus schlegeli	90.4	90.2	90.5	90.1	90.3	91.9	91.7	92.1	92.2	92.3	92.3	94.5	94.5	94.4	96.9	96.9	97	-	30	407	401	394	426
Eudyptes chrysolophus chrysolophus	90.4	90.2	90.4	90.1	90.2	91.9	91.7	92.1	92.2	92.2	92.3	94.5	94.5	94.5	96.9	97	97	99.8	-	410	397	396	430
Eudyptes sclateri	90.3	90.4	90.4	90.3	90.4	91.8	91.6	92.1	92.2	92.2	92.3	94.8	94.7	94.7	96.8	96.9	97	97.1	97.1	-	329	335	268
Eudyptes pachyrhynchus	90.3	90.2	90.3	90.2	90.2	91.7	91.7	92.2	92.3	92.3	92.3	94.6	94.6	94.6	97	97	97.1	97.2	97.2	97.7	-	110	337
Eudyptes robustus	90.3	90.3	90.4	90.4	90.4	91.8	91.8	92.3	92.4	92.4	92.4	94.7	94.7	94.6	97	97	97.2	97.2	97.2	97.6	99.2	-	339
Eudyptes warhami	90.3	90.3	90.4	90.4	90.4	91.8	91.7	92.1	92.3	92.3	92.3	94.7	94.7	94.7	96.9	96.9	97	97	97	98.1	97.6	97.6	-

Supplementary Table 7. Bone measurements (in mm) across all *Eudyptes* taxa. Asterisks denote estimates that are based on measurements from partially preserved bones. Note, the Mean, Standard Error (SE), and Number of Individuals (*n*) for each taxon's measurement is shown. With the exception of Museum Numbers labelled with 'CM', all samples were from NMNZ. 'X' indicates that the measurement was not taken from that individual, and 'NA' indicates that the Mean, SE, or *n* could not be obtained, as only one individual (or no individuals) were measured for that element.

Taxon	Museum Number	Age, sex	Skull (total length)	Internarial Bar (total length)	Premaxilla (maximum width at the caudal end of nares)	Internarial Bar (maximum width)	Skull (occipital condyle to nasofrontal flexure)	Skull (maximum width)	Mandible (total length)	Mandible (depth)	Coracoid (total length)	Humerus (total length)	Carpometacarpus (total length)	Tibiotarsus (distal width)
Eudyptes warhami	S.26908	Unknown	135.8*	73.5*	20.5	9.3	62.3	51.9	X	X	X	X	X	X
Eudyptes warhami	S.30440	Unknown	X	X	X	X	X	X	118.6	20.2	X	X	X	X
Eudyptes warhami	S.33007	Unknown	134.8	75	22.4*	8.3	59.8	52.6	X	X	X	X	X	X
Eudyptes warhami	S.24277	Unknown	X	X	X	X	X	X	X	X	X	X	43.2	X
Eudyptes warhami	S.25157	Unknown	X	X	X	X	X	X	X	X	X	69.1	X	X
Eudyptes warhami	S.27259	Unknown	X	X	X	X	X	X	X	X	79.4	X	X	X
Eudyptes warhami	S.47917	Unknown	X	X	X	X	X	X	X	X	80.3	X	X	X
Eudyptes warhami	S.47921	Unknown	X	X	X	X	X	X	X	X	X	X	X	12
Eudyptes warhami	CM Av.6816	Unknown	130*	X	X	8.9	60.56	54.3	X	X	X	X	X	X
Eudyptes warhami	CM Av.27407	Unknown	X	X	X	X	X	X	X	X	X	67	X	X
	Mean		**133.5**	**74.3**	**21.5**	**8.8**	**60.9**	**52.9**	**118.6**	**20.2**	**79.9**	**68.1**	**43.2**	**12**
	SE		**1.8**	**0.8**	**0.9**	**0.3**	**0.7**	**0.7**	**NA**	**NA**	**0.4**	**1.5**	**NA**	**NA**
	n		**3**	**2**	**2**	**3**	**3**	**3**	**1**	**1**	**2**	**2**	**1**	**1**
Eudyptes chrysolophus schlegeli	DM.702-S	Unknown	X	X	X	7.9	57.8	52.1	X	X	X	X	X	X
Eudyptes chrysolophus schlegeli	OR.26959	Adult, Male	137.9	82.4	21.9	9	59.5	52.9	127.5	24.2	87.6	76.9	46.4	15.9
Eudyptes chrysolophus schlegeli	OR.27787	Immature, Female	128.9	75.4	19.2	6.9	56.6	51.3	119.2	20.8	89.3	76.5	45.7	14.8
Eudyptes chrysolophus schlegeli	OR.29963a	Unknown	X	X	X	X	X	X	X	X	89.2	X	X	X

Taxon	Museum Number	Age, sex	Skull (total length)	Internarial Bar (total length)	Premaxilla (maximum width at the caudal end of nares)	Internarial Bar (maximum width)	Skull (occipital condyle to nasofrontal flexure)	Skull (maximum width)	Mandible (total length)	Mandible (depth)	Coracoid (total length)	Humerus (total length)	Carpometacarpus (total length)	Tibiotarsus (distal width)
Eudyptes chrysolophus chrysolophus (possibly *Eudyptes chrysolophus schlegeli*)	DM924-S	Immature, Female	X	X	X	X	X	X	X	X	88.3	X	X	X
Eudyptes chrysolophus chrysolophus (possibly *Eudyptes chrysolophus schlegeli*)	DM925-S	Unknown age, Male	X	X	X	X	X	X	X	X	92.6	X	X	X
Mean			**133.4**	**78.9**	**20.6**	**7.9**	**58**	**52.1**	**123.4**	**22.5**	**89.4**	**76.7**	**46.4**	**15.4**
SE			**4.5**	**3.5**	**1.4**	**0.6**	**0.8**	**0.5**	**4.2**	**1.7**	**0.9**	**0.2**	**0.3**	**0.6**
n			**2**	**2**	**2**	**3**	**3**	**3**	**2**	**2**	**5**	**2**	**2**	**2**
Eudyptes chrysolophus chrysolophus	OR.12841	Unknown age, Female	117.1	61.7	18.2	7.4	55.4	48.1	104	22.7	78.7	70.6	42.3	14.9
Eudyptes chrysolophus chrysolophus	OR.12840	Unknown age, Female	120.4	62.1	20.9	7.7	58.3	51.2	110.3	20.8	83.9	74.6	45	15.3
Eudyptes chrysolophus chrysolophus	DM703-S	Unknown	120.9	66	20.9	7.7	54.9	49.9	X	X	X	X	X	X
Mean			**119.5**	**63.3**	**20**	**7.6**	**56.2**	**49.7**	**107.2**	**21.8**	**81.3**	**72.6**	**43.7**	**15.1**
SE			**1.2**	**1.4**	**0.9**	**0.1**	**1.1**	**0.9**	**3.2**	**0.9**	**2.6**	**2**	**1.4**	**0.2**
n			**3**	**3**	**3**	**3**	**3**	**3**	**2**	**2**	**2**	**2**	**2**	**2**
Eudyptes filholi	OR.23674	Unknown	98.1	49.2	16.2	5.4	48.9	X	88.7	15.4	67.9	57.4	34.7	11.6
Eudyptes filholi	OR.19303	Unknown	112.4	59.7	19.9	7	52.7	47.4	102.6	19.1	72	61.8	37	11.8
Eudyptes filholi	OR.19291	Unknown age, Male	106	54	18.7	6.8	52	49.2	96.4	20.1	74.4	63.4	38.9	12.4
Eudyptes filholi	OR.19306	Unknown	105.8	53.1	18.4	6.6	52.7	44.9	97.6	16.6	72.6	61.7	37.7	12.5

Taxon	Museum Number	Age, sex	Skull (total length)	Internarial Bar (total length)	Premaxilla (maximum width at the caudal end of nares)	Internarial Bar (maximum width)	Skull (occipital condyle to nasofrontal flexure)	Skull (maximum width)	Mandible (total length)	Mandible (depth)	Coracoid (total length)	Humerus (total length)	Carpometacarpus (total length)	Tibiotarsus (distal width)
Eudyptes filholi	OR.9147	Unknown age, Female	105.1	54.3	17.1	6.2	50.8	44.4	95.3	17.7	69.1	58.5	35.5	12.4
Eudyptes filholi	DM624-S	Unknown	106.5	54.8	18.9	6.7	51.7	46.4	96.4	18.7	69	59.5	35.7	13
Eudyptes filholi	OR.11221	Unknown	112.1	58	20	7.2	54.1	48.9	100.1	18.3	71.7	62.7	39.1	12.6
Eudyptes filholi	OR.11236	Unknown	110.4	55.8	17.9	6.4	54.6	47.2	98.1	19.9	72.1	61.2	36.4	13
Mean			**107.1**	**54.9**	**18.4**	**6.5**	**52.2**	**46.9**	**96.9**	**18.2**	**71.1**	**60.8**	**36.9**	**12.4**
SE			**1.6**	**1.1**	**0.5**	**0.2**	**0.6**	**0.7**	**1.4**	**0.6**	**0.8**	**0.7**	**0.6**	**0.2**
n			**8**	**8**	**8**	**8**	**8**	**7**	**8**	**8**	**8**	**8**	**8**	**8**
Eudyptes moseleyi	S.34722.4	Unknown	X	X	16.9	6.6	X	X	X	X	X	X	X	X
Eudyptes moseleyi	S.34722.6	Unknown	X	X	X	6.4	X	X	X	X	X	X	X	X
Eudyptes moseleyi	S.34722.1	Unknown	X	X	X	X	51.3	46.3	X	X	X	X	X	X
Eudyptes moseleyi	S.34722.2	Unknown	X	X	X	X	54.6	X	X	X	X	X	X	X
Eudyptes moseleyi	S.34722.5	Unknown	X	X	X	X	X	X	94.9	16.9	X	X	X	X
Eudyptes moseleyi	S.34722.3	Unknown	X	X	X	X	X	X	X	17.4	X	X	X	X
Eudyptes moseleyi	S.34722.7	Unknown	X	X	X	X	X	X	X	X	X	59.1	X	X
Eudyptes moseleyi	S.34722.8	Unknown	X	X	X	X	X	X	X	X	X	X	X	12.1
Eudyptes moseleyi	S.34590.1	Unknown	X	X	X	X	X	X	X	X	X	59	X	X
Eudyptes moseleyi	S.34590.7	Unknown	X	X	X	X	X	X	X	X	X	X	34.2	X
Eudyptes moseleyi	S.34590.2	Unknown	X	X	X	X	X	X	X	X	X	X	X	12.2
Eudyptes moseleyi	S.34590.3	Unknown	X	X	X	X	X	X	X	X	X	X	X	12.1
Eudyptes moseleyi	S.34590.4	Unknown	X	X	X	X	X	X	X	X	X	X	X	12.5
Eudyptes moseleyi	S.34590.5	Unknown	X	X	X	X	X	X	X	X	X	X	X	12.7
Eudyptes moseleyi	S.34590.6	Unknown	X	X	X	X	X	X	X	X	X	X	X	12.5
Eudyptes moseleyi	S.34783	Unknown	X	X	X	X	X	X	X	X	X	X	X	12.1

Taxon	Museum Number	Age, sex	Skull (total length)	Internarial Bar (total length)	Premaxilla (maximum width at the caudal end of nares)	Internarial Bar (maximum width)	Skull (occipital condyle to nasofrontal flexure)	Skull (maximum width)	Mandible (total length)	Mandible (depth)	Coracoid (total length)	Humerus (total length)	Carpometacarpus (total length)	Tibiotarsus (distal width)
Eudyptes moseleyi	S.34711	Unknown	X	X	X	X	X	X	X	X	70.7	X	X	12.5
	Mean		**NA**	**NA**	**16.9**	**6.5**	**53**	**46.3**	**94.9**	**17.2**	**70.7**	**59.1**	**34.2**	**12.3**
	SE		**NA**	**NA**	**NA**	**0.1**	**1.7**	**NA**	**NA**	**0.3**	**NA**	**0.1**	**NA**	**0.1**
	n		**NA**	**NA**	**1**	**2**	**2**	**1**	**1**	**2**	**1**	**2**	**1**	**8**
Eudyptes sclateri	OR.1441	Unknown	120.5	63.7	21.1	7.7	56.8	52.7	X	24.9	85	72	43.4	15.6
Eudyptes sclateri	OR.11217	Unknown	121.7	64.9	20.4	7.4	56.8	51.5	114.2	24.8	83	73.1	45.5	15
Eudyptes sclateri	OR.15168	Immature, Male	117.1	59.6	19.4	7.2	57.5	50.4	107.2	23.8	84.8	72.8	45.6	15.4
Eudyptes sclateri	OR.18897	Unknown	115.9	60.8	18.6	6.9	55.1	50.5	108.2	23.7	79	70	43.6	14.7
Eudyptes sclateri	DM669-S	Unknown	124.2	65.3	19.6	7.8	58.9	52.9	116.1	25.7	88.7	75.3	47	16.2
Eudyptes sclateri	DM670-S	Unknown	125.1	69.6	20.9	6.7	55.5	52.9	121.6	25	88.8	78	49.1	15.4
Eudyptes sclateri	DM672-S	Unknown	118.2	60.3	19.5	7.2	57.9	51.2	107.9	23	83.8	72.7	45.7	15.7
Eudyptes sclateri	DM668-S	Unknown	129.7	72.1	20.8	7.1	57.6	53	123.7	26.5	90	74.5	47	15.9
Eudyptes sclateri	OR.23160	Adult, Male	123	66.5	20.3	7.4	56.5	52.4	117.4	25.1	89.7	75.7	46	17
Eudyptes sclateri	OR.23578	Unknown	127.1	69.9	X	6.8	57.2	51.6	117.2	25.2	88.5	74	47.5	15.1
Eudyptes sclateri	OR.25560	Unknown age, Male	123.1	64.5	19.8	7.4	58.6	49.9	113.6	24.4	84.8	73.6	47.2	15.4
Eudyptes sclateri	OR.25674	Unknown	119.3	63.2	19.6	7.1	56.1	52	110.1	24.3	85.8	77.2	45	15.5
Eudyptes sclateri	OR.25558	Female	115.6	58.5	17.5	7.5	57.1	49	106.3	22.5	79.4	69.8	42.5	14.3
	Mean		**121.6**	**64.5**	**19.8**	**7.2**	**57**	**51.5**	**113.6**	**24.5**	**85.5**	**73.7**	**45.8**	**15.5**
	SE		**1.2**	**1.2**	**0.3**	**0.1**	**0.3**	**0.4**	**1.7**	**0.3**	**1**	**0.7**	**0.5**	**0.2**
	n		**13**	**13**	**12**	**13**	**13**	**13**	**12**	**13**	**13**	**13**	**13**	**13**
Eudyptes robustus	OR.23754	Unknown	117	60.7	22	6.6	56.3	50.2	107.6	18.4	74.1	63.6	40	12.7
Eudyptes robustus	OR.23755	Unknown	130.1	70	21.7	7.7	60.1	51.6	119	20.4	78.5	66.4	42.2	13.3
Eudyptes robustus	OR.23735	Unknown	115.2	59.3	21.1	7.1	55.9	50.4	104.6	17.9	71.5	62.8	38.9	12.7
Eudyptes robustus	OR.23736	Adult, Male	125.5	69.3	23.8	8.5	56.2	53.6	115.6	19.7	77.1	67.2	41	13.6
Eudyptes robustus	OR.23737	Unknown age, Female	113	56.4	22.1	7.2	56.6	50.7	99.2	18	71.4	61	37.6	12.8

Taxon	Museum Number	Age, sex	Skull (total length)	Internarial Bar (total length)	Premaxilla (maximum width at the caudal end of nares)	Internarial Bar (maximum width)	Skull (occipital condyle to nasofrontal flexure)	Skull (maximum width)	Mandible (total length)	Mandible (depth)	Coracoid (total length)	Humerus (total length)	Carpometacarpus (total length)	Tibiotarsus (distal width)
Eudyptes robustus	OR.23741	Unknown	125.4	68.1	22.6	7.5	57.3	53	118.5	19.9	75.2	68.3	42.6	14.2
Eudyptes robustus	OR.23746	Unknown	127.1	69.2	23.8	8	57.9	52.8	118.5	20.2	72.2	63.4	41.6	14.1
Eudyptes robustus	OR.23758	Unknown age, Male	128.3	69.2	22.5	8	59.1	53	119.9	20.7	76.3	68.2	41.9	13.9
Eudyptes robustus	OR.26826	Unknown	123.4	64.6	22.2	8.4	58.8	51.1	112.4	19.5	73.8	65.8	40.7	13.5
Eudyptes robustus	OR.23672	Immature, Female	106.2	50	19.8	6.5	56.2	49.6	104	18.2	73.5	63.3	39.9	12.8
Eudyptes robustus	OR.27147a	Adult, Unknown sex	131.7	72.3	24.1	8	59.4	53.5	122.5	20.6	74.7	66.8	39.4	13.7
Eudyptes robustus	OR.23673	Immature Female	114.6	60.6	20.4	6.7	54	49.3	102.6	17.7	71.2	62.4	38.1	13.6
Eudyptes robustus	OR.23724	Unknown	126.4	69.6	21.1	7.4	56.8	49.6	113.5	19.5	73.9	66	41	13.3
Eudyptes robustus	OR.30196	Unknown	124.8	65.9	25.7	8.4	58.9	54	114.9	19.6	77.2	68.2	42.3	14.4
	Mean		**122.1**	**64.7**	**22.4**	**7.6**	**57.4**	**51.6**	**112.3**	**19.3**	**74.3**	**65.2**	**40.5**	**13.5**
	SE		**2**	**1.7**	**0.4**	**0.2**	**0.5**	**0.4**	**2**	**0.3**	**0.6**	**0.7**	**0.4**	**0.2**
	n		**14**	**14**	**14**	**14**	**14**	**14**	**14**	**14**	**14**	**14**	**14**	**14**
Eudyptes pachyrhynchus	OR.11230	Unknown	112.1	56.7	20	7.4	55.4	49.6	102.5	18.1	76	65.9	39.2	14.2
Eudyptes pachyrhynchus	OR.13308	Unknown	120.8	63.3	21.2	8.1	57.5	51	111.6	20	76.7	65.8	39.2	14.2
Eudyptes pachyrhynchus	OR.24426	Adult, Male	118.3	61	21.2	7.4	57.3	52.6	109.2	18	73.3	66	40.7	14.5
Eudyptes pachyrhynchus	OR.24427	Adult, Male	119.6	62.6	22.7	7.9	57	54.3	109.1	19.8	74.7	66.7	40.9	15
Eudyptes pachyrhynchus	OR.24428	Adult, Female	113.8	57.6	19.8	6.6	56.2	49.5	103.9	17.1	74.1	65.4	40.8	14
Eudyptes pachyrhynchus	OR.29650	Unknown age, Female	111.5	54.7	21.2	7.2	56.8	51.3	100.8	18.1	73.2	65.1	38.9	13.7
Eudyptes pachyrhynchus	OR.24546	Adult, Female	114.6	56.8	21.5	8.1	57.8	51.6	105.4	19.6	74.4	65.3	40.4	14.4
Eudyptes pachyrhynchus	OR.24429	Unknown	108.8	54.4	18.2	6.5	54.4	46.9	97.8	18.3	72.9	62.9	38	13.6
Eudyptes pachyrhynchus	OR.24514	Unknown age, Male	118.1	61.6	21.8	7.2	56.5	53.3	109.3	19	77.6	68	41.2	14.2
Eudyptes pachyrhynchus	OR.29077	Adult, Female	109.2	54.3	20.9	7.3	54.9	48.9	101.3	17.7	72	64.8	39	14.1

Taxon	Museum Number	Age, sex	Skull (total length)	Internarial Bar (total length)	Premaxilla (maximum width at the caudal end of nares)	Internarial Bar (maximum width)	Skull (occipital condyle to nasofrontal flexure)	Skull (maximum width)	Mandible (total length)	Mandible (depth)	Coracoid (total length)	Humerus (total length)	Carpometacarpus (total length)	Tibiotarsus (distal width)
Eudyptes pachyrhynchus	OR.29154	Adult, Male	117.7	58.9	21.5	7.6	58.8	51.3	109	19.3	74.6	67.4	41.9	14.4
Eudyptes pachyrhynchus	OR.30090	Adult, Female	112.1	57.1	20.6	6.9	55	49.5	99.9	17	71.6	64.5	36.5	13.9
Mean			**114.7**	**58.3**	**20.9**	**7.4**	**56.5**	**50.8**	**105**	**18.5**	**74.3**	**65.7**	**39.7**	**14.2**
SE			**1.2**	**0.9**	**0.3**	**0.2**	**0.4**	**0.6**	**1.3**	**0.3**	**0.5**	**0.4**	**0.4**	**0.1**
n			**12**	**12**	**12**	**12**	**12**	**12**	**12**	**12**	**12**	**12**	**12**	**12**

Supplementary Table 8. Comparative measurements (in mm) across all *Megadyptes* taxa, as summarised from *Supplementary Tables 9 – 12*. Where two sets of measurements are presented, the first set of measurements are from Boessenkool *et al*. (2008) and the second measurement is from either *Supplementary Table 9* or *Supplementary Table 10*. 'NA' indicates that the Mean, SE, or Number of Individuals (*n*) could not be obtained, as only one individual (or no individuals) were measured for that element.

Element	*Megadyptes antipodes antipodes*			*Megadyptes antipodes waitaha*			*Megadyptes antipodes richdalei*		
	Mean	**SE**	***n***	**Mean**	**SE**	***n***	**Mean**	**SE**	***n***
Skull (total length)	132.6	1.5	8	NA	NA	NA	115.8	NA	1
Premaxilla (maximum width)	6.3	0.2	8	NA	NA	NA	5.4	0	2
Mandible (total length)	126.3	1.8	9	NA	NA	NA	116.9	1.2	2
Mandible (depth)	15.5	.03	9	NA	NA	NA	15.0	0.9	2
Coracoid (total length)	89.1/91.0	0.5/0.8	26/9	81.1/79.6	0.5/1.1	11/4	78.9	3	2
Coracoid (proximal width)	29.8	0.3	9	24.4	0.7	3	23.8	0	2
Humerus (total length)	78.3/77.9	0.4/0.4	26/9	72.2/70.6	0.5/0.9	25/9	67.5	1.1	3
Humerus (proximal width)	23.1	0.3	9	20.9	0.2	11	19.9	0.4	3
Humerus (distal width)	21.8	0.2	9	19.6	0.3	10	20.2	0.7	3
Humerus (mid-shaft width)	13.5/16.5	0.1/0.2	26/9	12.6/14.5	0.1/0.3	25/11	13.6	1.4	3
Ulna (total length)	60.2	0.5	9	56.2	NA	1	53.9	NA	1
Radius (total length)	58.4	0.5	9	53.5	0.7	3	52.7	NA	1
Carpometacarpus (total length)	46.1	0.4	9	43.7	NA	1	41.7	NA	1
Femur (total length)	85.2/85.1	0.4/0.9	33/9	77.3/75.1	0.5/0.4	28/10	73.4	NA	1
Femur (mid-shaft width)	8.9/9.0	0.1/0.1	33/9	8.3/7.8	0.1/0.1	28/19	8.2	NA	1
Femur (proximal width)	18.4/19.2	0.2/0.3	31/9	16.4/16.5	0.2/0.2	18/10	17.0	NA	1
Femur (distal width)	19.7/18.4	0.2/0.2	32/9	17.8/15.8	0.2/0.2	22/12	15.4	NA	1
Tibiotarsus (total length)	131.8	1.1	9	NA	NA	NA	115.9	NA	1
Tibiotarsus (proximal width)	13.8	0.2	9	11.7	NA	1	12.2	0.1	2
Tibiotarsus (distal width)	16.8	0.2	9	13.8	0.7	3	15.4	NA	1
Tarsometatarsus (total length)	35.7/35.3	0.2/0.3	26/9	33.4/31.6	0.3/0.4	11/15	31.5	NA	1
Tarsometatarsus (proximal width)	18.8/18.4	0.1/0.2	26/9	17.3/16.4	0.3/0.3	10/14	16.7	NA	1
Tarsometatarsus (distal width)	23.1/23.0	0.2/0.3	26/9	21.6/20.5	0.4/0.3	11/15	21.1	NA	1

Supplementary Table 9. Measurements (in mm) of *Megadyptes* bones derived from the Chatham Islands. Samples numbers with 'NMNZ' are from The National Museum of New Zealand Te Papa Tongarewa, 'AM' are from the Auckland Museum and 'CM' are from the Canterbury Museum. 'X' indicates that the measurement was not taken from that individual.

Element	NMNZ S.26921	NMNZ S.47918	NMNZ S.30968	NMNZ S.47765	NMNZ S.45876	CM Av11287	CM (ACAD12997)	AM LB12063	Mean
Skull (total length)	X	X	X	X	115.8	X	X	X	115.8
Premaxilla (maximum width)	X	X	X	5.4	5.4	X	X	X	5.4
Mandible (total length)	115.7	X	118.0	X	X	X	X	X	116.9
Mandible (depth)	14.1	X	15.9	X	X	X	X	X	15.0
Coracoid (total length)	81.9	75.9	X	X	X	X	X	X	78.9
Coracoid (proximal width)	23.8	23.8	X	X	X	X	X	X	23.8
Humerus (total length)	69.4	X	X	X	X	67.5	65.5	X	67.5
Humerus (proximal width)	20.6	X	X	X	X	19.5	19.5	X	19.9
Humerus (distal width)	19.1	X	X	X	X	21.5	20	X	20.2
Humerus (mid-shaft width)	14.3	X	X	X	X	14.5	12	X	13.6
Ulna (total length)	53.9	X	X	X	X	X	X	X	53.9
Radius (total length)	52.7	X	X	X	X	X	X	X	52.7
Carpometacarpus (total length)	41.7	X	X	X	X	X	X	X	41.7
Femur (total length)	73.4	X	X	X	X	X	X	X	73.4
Femur (mid-shaft width)	8.2	X	X	X	X	X	X	X	8.2
Femur (proximal width)	17.0	X	X	X	X	X	X	X	17.9
Femur (distal width)	15.4	X	X	X	X	X	X	X	15.4
Tibiotarsus (total length)	115.9	X	X	X	X	X	X	X	115.9
Tibiotarsus (proximal width)	12.3	X	X	X	X	X	X	12.1	12.2
Tibiotarsus (distal width)	15.4	X	X	X	X	X	X	X	15.4
Tarsometatarsus (total length)	31.5	X	X	X	X	X	X	X	31.5
Tarsometatarsus (proximal width)	16.7	X	X	X	X	X	X	X	16.7
Tarsometatarsus (distal width)	21.1	X	X	X	X	X	X	X	21.1

Supplementary Table 10. Measurements (in mm) of *Megadyptes antipodes waitaha* bones. Some specimen numbers represent multiple individuals that were found in association. The mid-shaft width of the humerus is measured level with the preaxial angle. The tibiotarsus proximal width measurement excludes the cnemial crest. All samples are from NMNZ. 'X' indicates that the measurement was not taken from that sample.

Element	S.35426	S.35901	S.35913	S.36277	S.36370	S.36744	S.38489	S.38496	S.38513	S.38525	S.38530	S.38564	S.38949	S.38966	S.42156	S.43867	Mean
Coracoid (total length)	X	X	X	82.4	78.3, 77.5	X	X	X	80.2	X	X	X	X	X	X	X	79.6
Coracoid (proximal width)	X	X	X	23.6	25.8	X	X	X	23.7	X	X	X	X	X	X	X	24.4
Humerus (total length)	65.9	X	X	71.0, 74.7, 73.0, 69.9, 73.3	69.1, 68.1, 70.1	73.4	X	X	X	X	X	X	X	X	X	X	70.6
Humerus (proximal width)	21.1, 18.9	X	20.7	21.0, 21.6, 21.6, 20.9, 20.3	20.4, 21.1, 21.1	21.4	X	X	X	X	X	X	X	X	X	X	20.9
Humerus (distal width)	19.3	X	X	20.3, 21.4, 20.4, 19.2, 18.8	19.2, 19.4, 19.3	18.9	X	X	X	X	X	X	18.7	X	X	X	19.6
Humerus (mid-shaft width)	13.5, 13.9, 12.6	X	X	15.9, 15.1, 15.4, 15.0, 15.2	14.6, 14.5, 15.3	13.8	X	X	X	X	X	X	X	X	X	X	14.5
Ulna (total length)	X	X	56.2	X	X	X	X	X	X	X	X	X	X	X	X	X	56.2
Radius (total length)	X	X	X	X	X	X	X	X	54.5	52.2	X	X	X	X	53.7	X	53.5
Carpometacarpus (total length)	X	X	43.7	X	X	X	X	X	X	X	X	X	X	X	X	X	43.7
Femur (total length)	74.5, 75.5	X	X	X	74.6, 74.4	X	X	77.1, 74.8	X	X	73.8	X	77.7	73.0	75.1	X	75.1
Femur (mid-shaft width)	8.0, 7.3	X	8.0	X	8.0, 7.9	X	X	7.7, 7.8, 8.8	X	X	8.1, 7.6, 8.4	6.5, 7.7	7.7	7.3, 7.1	7.7, 7.6	8.2	7.8
Femur (proximal width)	17.2, 16.4	X	15.3, 16.8	X	17.2, 17.2	X	X	16.8	X	X	X	15.8	15.7	16.4, 15.9	16.1	X	16.5
Femur (distal width)	17.0, 15.0	X	16.2	X	16.4, 16.2	X	X	15.2	X	X	15.0	X	15.5	14.9, 16.0	15.8, 16.2	X	15.8
Tibiotarsus (proximal width)	11.7	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	11.7
Tibiotarsus (distal width)	15.1	X	X	X	X	X	X	X	X	X	X	X	X	12.7	13.7	X	13.8

Element	S.35426	S.35901	S.35913	S.36277	S.36370	S.36744	S.38489	S.38496	S.38513	S.38525	S.38530	S.38564	S.38949	S.38966	S.42156	S.43867	Mean
Tarsometatarsus (total length)	34.0	X	31.9, 29.6	X	31.7, 32.0, 31.0, 28.4, 31.6, 29.9	X	32.1	30.0, 32.0	X	X	X	32.3	X	33.3	34.2	X	31.6
Tarsometatarsus (proximal width)	17.8	18.6	17.0, 16.3	X	15.6, 15.8, 15.0, 14.4, 16.5, 15.2	X	X	16.9, 16.5	X	X	X	X	X	16.8	16.9	X	16.4
Tarsometatarsus (distal width)	22.5	X	22.0, 20.4	X	20.4, 19.2, 19.7, 19.3, 21.0, 20.0	X	18.6	20.1, 21.5	X	X	X	20.3	X	20.9	21.3	X	20.5

Supplementary Table 11. Measurements (in mm) of *Megadyptes antipodes antipodes* and *M. a. waitaha* bones, derived from Table S3 in Boessenkool *et al*. (2008). Note that Boessenkool *et al*. (2008) did not describe the measurement methods for the humerus width, so the measurements included in *Supplementary Tables 7 – 10* and *Supplementary Table 12* may not be directly comparable. Note also that Boessenkool *et al*. (2008) apparently transposed the distal and proximal width measurements for the tarsometatarsus, which we have corrected below.

Bone type	Measurement	*Megadyptes antipodes antipodes*			*Megadyptes antipodes waitaha*		
		n	**Mean**	**SE**	***n***	**Mean**	**SE**
Femur	Length	33	85.2	0.4	28	77.3	0.5
	Shaft width	33	8.9	0.1	28	8.3	0.1
	Distal width	32	19.7	0.2	22	17.8	0.2
	Proximal width	31	18.4	0.2	18	16.4	0.2
Humerus	Length	26	78.3	0.4	25	72.2	0.5
	Width	26	13.5	0.1	25	12.6	0.1
Coracoid	Length	26	89.1	0.5	11	81.1	0.5
Tarsometatarsus	Length	26	35.7	0.2	11	33.4	0.3
	Proximal width	26	18.8	0.1	10	17.3	0.3
	Distal width	26	23.1	0.2	11	21.6	0.4

Supplementary Table 12. Measurements (in mm) of modern *Megadyptes antipodes antipodes* bones. Only individuals with well-ossified skeletons were measured (e.g. juveniles were excluded). The mid-shaft width of the humerus was measured level with the preaxial angle, and the tibiotarsus proximal width measurement excludes the cnemial crest. All samples are from NMNZ. 'X' indicates that the measurement was not taken from that sample.

Element	OR.24638	OR.24690	OR.26794	OR.26470	OR.19308	OR.18335	DM.595	OR.23135	DM.631	Mean
Age, Sex	Adult, Male	Adult, Male	Adult, Male	Adult, Unknown sex	Unknown	Unknown	Unknown	Adult, Female	Unknown	NA
Skull (total length)	132.1	X	135.3	137.2	133.1	133.2	137.1	124.6	128.2	132.6
Premaxilla (maximum width)	6.5	6.3	6.9	6.6	6.6	6.4	6.5	5.2	6.1	6.3
Mandible (total length)	128.0	126.8	130.5	131.7	125.0	127.0	131.5	115.1	121.4	126.3
Mandible (depth)	16.1	15.4	16.6	15.7	14.1	15.6	16.4	14.7	15.1	15.5
Coracoid (total length)	91.0	88.6	92.3	93.0	89.7	88.7	95.5	88.3	91.9	91.0
Coracoid (proximal width)	29.5	29.0	30.5	29.3	30.0	29.8	31.4	30.7	28.0	29.8
Humerus (total length)	77.5	77.0	79.2	78.8	77.2	76.4	80.2	76.8	77.9	77.9
Humerus (proximal width)	23.2	23.2	24.0	23.8	23.8	22.1	23.8	22.2	21.8	23.1
Humerus (distal width)	21.0	22.0	22.3	22.8	22.1	21.2	22.4	20.9	21.2	21.8
Humerus (mid-shaft width)	16.3	16.5	16.9	17.3	16.2	16.3	17.6	16.0	15.8	16.5
Ulna (total length)	60.5	60.1	61.5	61.1	59.3	57.4	63.1	59.1	60.1	60.2
Radius (total length)	59.6	58.4	59.1	59.3	57.0	56.0	61.0	57.2	58.3	58.4
Carpometacarpus (total length)	46.8	46.2	47.4	47.3	45.4	43.8	47.5	44.5	45.6	46.1
Femur (total length)	85.4	83.8	86.4	87.6	84.2	82.5	90.1	82.5	83.2	85.1
Femur (mid-shaft width)	8.8	9.0	9.2	9.4	9.1	8.7	9.5	8.6	8.7	9.0
Femur (proximal width)	19.0	19.3	19.7	20.2	20.1	17.8	19.9	18.9	18.3	19.2
Femur (distal width)	18.6	18.5	18.6	18.3	18.6	18.1	19.3	18.4	17.5	18.4
Tibiotarsus (total length)	133.5	129.5	135.0	133.9	129.7	128.8	137.6	128.3	129.9	131.8
Tibiotarsus (proximal width)	14.2	13.3	13.8	14.6	14.3	12.9	14.4	13.5	13.6	13.8
Tibiotarsus (distal width)	17.1	17.1	17.3	17.2	16.3	16.1	17.7	16.4	16.0	16.8
Tarsometatarsus (total length)	35.5	35.8	36.0	36.1	34.3	34.2	36.6	34.9	34.7	35.3
Tarsometatarsus (proximal width)	18.5	18.6	18.7	18.9	17.9	18.0	19.0	18.5	17.5	18.4

Supplementary Table 13. Divergence-dates (Ma) between all penguin taxa based on *Figure 6B*. Divergence dates are based on near-complete mitogenomes, using the birth-death speciation prior.

Taxon 1	Taxon 2	Upper 95 % HPD	Lower 95 % HPD	Geological Period
Aptenodytes, Pygoscelis	*Eudyptula, Spheniscus, Eudyptes, Megadyptes*	22	12.5	Miocene
Aptenodytes	*Pygoscelis*	20.4	10.2	Miocene
Megadyptes, Eudyptes	*Eudyptula, Spheniscus*	17.1	10.6	Miocene
Eudyptula	*Spheniscus*	13.9	9.3	Miocene
Pygoscelis antarctica, P. papua	*Pygoscelis adeliae*	10.4	4.9	Pliocene
Megadyptes	*Eudyptes*	9.3	4.7	Pliocene
Pygoscelis papua	*Pygoscelis antarctica*	8.2	3.5	Pliocene
Eudyptes chrysolophus chrysolophus, E. chrysolophus schlegeli	*Eudyptes pachyrhynchus, E. robustus, E. sclateri, E. warhami, E. moseleyi, E. filholi, E. chrysocome*	5.3	2.7	Pliocene
Aptenodytes patagonicus	*Aptenodytes forsteri*	5.1	2.0	Pliocene
Eudyptes moseleyi, E. filholi, E. chrysocome	*Eudyptes pachyrhynchus, E. robustus, E. sclateri, E. warhami*	4.9	2.4	Pliocene
Eudyptula novaehollandiae	*Eudyptula minor*	4.6	2.0	Pliocene
Eudyptes pachyrhynchus, E. robustus	*Eudyptes sclateri, E. warhami*	3.5	1.7	Pleistocene
Spheniscus humboldti, S. mendiculus	*Spheniscus demersus, S. magellanicus*	3.0	1.4	Pleistocene
Eudyptes chrysocome, E. filholi	*Eudyptes moseleyi*	2.7	1.2	Pleistocene
Eudyptes sclateri	*Eudyptes warhami*	2.5	1.1	Pleistocene
Spheniscus demersus	*Spheniscus magellanicus*	2.2	0.9	Pleistocene
Spheniscus humboldti	*Spheniscus mendiculus*	1.6	0.6	Pleistocene
Eudyptes pachyrhynchus	*Eudyptes robustus*	1.4	0.5	Pleistocene
Eudyptes chrysocome	*Eudyptes filholi*	1.3	0.5	Pleistocene
Megadyptes antipodes waitaha	*Megadyptes antipodes antipodes, Megadyptes antipodes richdalei*	0.7	0.2	Pleistocene
Megadyptes antipodes antipodes	*Megadyptes antipodes richdalei*	0.4	0.1	Pleistocene
Eudyptes chrysolophus chrysolophus	*Eudyptes chrysolophus schlegeli*	0.2	0.0	Pleistocene

Chapter 3

Receding ice drove parallel expansions in eight Southern Ocean penguin taxa

Publication details, contributions and acknowledgements

Parts of this chapter are under revision in:

Theresa L Cole[1,2], Ludovic Dutoit[1], Nic Dussex[3,4], Tom Hart[3], Alana Alexander[4], Jane L Younger[5], Gemma V Clucas[6,7], María José Frugone[8,9],Yves Cherel[11], Richard Cuthbert[12,13], Ursula Ellenberg[14,15], Steven R Fiddaman[16], Johanna Hiscock[17], David M Houston[18], Pierre Jouventin[19], Thomas Mattern[1], Gary Miller[20,21], Colin M Miskelly[22], Paul Nolan[23], Michael J. Polito[24], Petra Quillfeldt[25], Peter G Ryan[26], Adrian Smith[16], Alan JD Tennyson[22], David R Thompson[27], Barbara Wienecke[28], Juliana A Vianna[9] & Jonathan M Waters[1]. (under revision). Receding ice drove rapid expansions in eight Southern Ocean penguins. *Proceedings of the National Academy of Sciences.*

Author affiliations:

[1]Department of Zoology, University of Otago, PO Box 56, Dunedin 9054, Otago New Zealand. [2]Manaaki Whenua Landcare Research, PO Box 69040, Lincoln 7640, Canterbury, New Zealand. [3]Department of Bioinformatics and Genetics, Swedish Museum of Natural History, Box 50007, Stockholm 10405, Sweden. [4]Department of Anatomy, University of Otago, PO Box 56, Dunedin 9054, Otago, New Zealand. [5]Department of Zoology, University of Oxford, 11a Mansfield Road, Oxford OX1 3SZ, United Kingdom. [5]Milner Centre for Evolution, University of Bath, Claverton Down, Bath BA2 7AY, United Kingdom. [6]Atkinson Center for a Sustainable Future, Cornell University, Ithaca, New York 14850, United States of America. [7]Cornell Lab of Ornithology, Cornell University, Ithaca, New York 14850, United States of America. [8]Laboratorio de Ecología Molecular, Departamento de Ciencias Ecológicas II, Facultad de ciencias, Universidad de Chile, Las Encinas #3770, Ñuñoa 7750000, Santiago, Chile. [9]Pontificia Universidad Católica de Chile, Facultad de Agronomía e Ingeniería Forestal, Departamento de Ecosistemas y Medio Ambiente, Vicuña Mackena 4860, Macul 7810000, Santiago, Chile. [11]Centre d'Etudes Biologiques de Chizé, UMR 7372 du CNRS-La, Rochelle Université, Villiers-en-Bois 79360, France. [12]Royal Society for the Protection of Birds, The Lodge, Sandy, Bedfordshire SG19 2DL, United Kingdom. [13]World Land Trust, Blyth House, Halesworth, Suffolk IP19 8AB, United Kingdom. [14]Department of Ecology, Environment and Evolution, La Trobe University, Melbourne 3083, Victoria, Australia. [15]Global Penguin Society, University of Washington, Seattle 98195, United States of America. [16]Department of Zoology, University of Oxford, Peter Medawar Building for Pathogen Research, South Parks Road, Oxford OX1 3SY, United Kingdom. [17]Department of Conservation, Murihiku District Office, Invercargill, New Zealand. [18]Biodiversity, Department of Conservation, Private Bag 68908, Wellesley Street, Auckland 1141, Auckland New Zealand. [19]Centre d'Ecologie Fonctionnelle et Evolutive, UMR 5175, Campus CNRS, 1919 Route de Mende, 34293 Montpellier Cedex 5, France. [20]Division of Pathology and Laboratory Medicine, University of Western Australia, Crawley 6009, Western Australia, Australia. [21]Institute for Marine and Antarctic Studies, University of Tasmania, Hobart 7001, Tasmania, Australia. [22]Museum of New Zealand Te Papa Tongarewa, PO Box 467, Wellington 6140, Wellington,

New Zealand. [23]Department of Biology, The Citadel, 171 Moultrie St, Charleston 29409, South Carolina, United States of America. [24]Department of Oceanography and Coastal Sciences, Louisiana State University, 1239 Energy, Coast and Environment Building, Baton Rouge 70803, Los Angeles, United States of America. [25]Justus Liebig Universität Gießen, Heinrich-Buff-Ring 26, Gießen 35392, Germany. [26]DST-NRF Centre of Excellence, FitzPatrick Institute of African Ornithology, University of Cape Town, Rondebosch 7701, Republic of South Africa. [27]National Institute of Water and Atmospheric Research Ltd., Private Bag 14901, Kilbirnie 6241, Wellington, New Zealand. [28]Department of the Environment and Energy, Australian Antarctic Division, 203 Channel Highway, Kingston 7050, Tasmania, Australia.

Author contributions:

Theresa L Cole (90%) and Jonathan M Waters (10%) conceived and designed the study. Theresa L Cole (98%) and Steven R Fiddaman (2%) undertook the laboratory work. Gemma V Clucas (45%), Jane L Younger (45%) and Tom Hart (10%) contributed published sequence data. Theresa L Cole (94%), Ludovic Dutoit (2%), Nicolas Dussex (2%), Jane L Younger (1%) and Gemma V Clucas (1%) analysed the genetic structure data. Ludovic Dutoit (30%), Nicolas Dussex (30%), Alana Alexander (30%) and Theresa L Cole (10%) undertook the demographic analyses. Peter G Ryan (8%), David R Thompson (8%), Tom Hart (8%), Yves Cherel (8%), Richard Cuthbert (5%), Gary Miller (5%), Paul Nolan (5%), Petra Quillfeldt (5%), Theresa L Cole (5%), Thomas Mattern (5%), Johanna Hiscock (5%), David M Houston (5%), Michael J Polito (5%), Adrian Smith (5%), Alan JD Tennyson (5%), Colin M Miskelly (5%), Barbara Wienecke (5%), Steven R Fiddaman (2%) and Pierre Jouventin (1%) collected blood samples. Theresa L Cole (60%), Tom Hart (30%) and Jonathan M Waters (10%) contributed to sequencing and laboratory costs. All co-authors helped to write the manuscript. Ludovic Dutoit, Nicolas Dussex, Tom Hart and Alana Alexander contributed equally and are joint second-author on the submitted article.

Acknowledgements:

We thank Catherine Bazjak, Anne-Sophie Coquel, Nina Dehnhard, Mike Fawcett, Holly Irvine, Kyle Morrison and Marion Nicolaus for sample collection, Tania King, Caroline Mitchell and Kat Trought for laboratory assistance, Vikram Chhatre, Arthur Georges, Bernd Gruber, Francisco Pina Martins, Ben Roberts, Albert Savary, Jamie Wood and Xander Xue for bioinformatics assistance and Luciano Beheregaray, Mike Hickerson, Andrzej Kilian, Michael Knapp, Paul Scofield, Paul Sunnucks, Janet Wilmshurst and Jamie Wood for discussions. We thank several anonymous reviewers for suggestions that greatly improved the manuscript. The research was approved by the Otago University Animal Ethics Committee 61/2016 (NZ), Oxford University Local Animal Ethics Committee (South Georgia, South Sandwich and Antarctic Peninsula), University of Western Australia Animal Ethics Committee (Macquarie Island), Woods Hole Oceanographic Institution Animal Care and Use Committee (Antarctic Peninsula), IPEV ethics committee (Amsterdam, Crozet and Kerguelen Islands, and Adélie Land), Ngāi Tahu Research Consultation Committee and The Zoological Society of London. Work was carried out under NZ Department of Conservation Permits (OT-25557-DOA, IACUC-18958.00, 32202-FAU, 35682-FAU, 37312-LND, 50437-DOA, 50436-FAU, 50464-DOA, and 38882-RES), NZ Ministry of Primary Industries (2015056535, 2016060908, 2017064905), a Permit to Possess Threatened Fauna for Scientific Purposes No. TFA 15086 (Macquarie Island), DPIWE Permit to Take Wildlife for Scientific Purposes No. FA05246 (Macquarie Island), Tristan Da Cunha Conservation Department (Gough Island), South African Department of Environmental Affairs, Government of South Georgia and the South Sandwich Islands Restricted Activity Permits (GSGSSI RAP 2018/018, GSGSSI), United States National Science Foundation Department of Polar Programs ACA permits (ACA 2016-011, ACA 2016-012), Falkland Islands Government (R05/2009, R014/2006) and United Kingdom Antarctic Permits. We thank Neil Fowke and Bruce McKinlay for NZ permit approval. This research was supported by Manaaki

Whenua Landcare Research, the University of Otago, NMNZ, United States National Science Foundation (OPP-012-8913), Quark Expeditions, Cheesemans Ecology Safaris, Golden Fleece Expeditions, New Island Conservation Trust, Deutsche Forschungsgemeinschaft, The Citadel Foundation, The Dalio Foundation, donations to Penguin Watch, the Institut Polaire Français Paul Emile Victor (IPEV, Programme N°109, Pierre Jouventin, Henri Weimerskirch, and N°134, Charlie A Bost), The Royal Society of NZ Hutton Fund, The Ornithological Society of NZ and an Alumni of Otago in America Award. Theresa L Cole was supported by an Otago University Postgraduate Award and an Otago University Postdoctoral Publishing Bursary.

Abstract

Climate shifts are key drivers of ecosystem change. Despite the critical importance of Antarctica and the Southern Ocean for global climate, the extent of climate-driven ecological change in this region remains controversial. In particular, the biological effects of rapidly changing sea-ice conditions remain poorly understood. We hypothesise that rapid post-glacial reductions in sea-ice drove biological shifts across multiple widespread Southern Ocean species. We test for demographic shifts driven by climate events over recent millennia by analysing population-genomic datasets spanning 12 Southern Ocean penguin taxa across three genera (*Eudyptes*, *Pygoscelis* and *Aptenodytes*). Demographic models for *Eudyptes filholi*, *E. chrysolophus chrysolophus*, *E. c. schlegeli*, *Aptenodytes patagonicus*, *A. forsteri*, *Pygoscelis adeliae*, *P. papua* and *P. antarcica* that inhabit Antarctic and sub-Antarctic coastlines affected by heavy sea-ice conditions during the LGM yielded genetic signatures of near-simultaneous population expansions that were likely associated with post-glacial warming. Of these eight taxa, populations of the ice-adapted *A. forsteri* began to expand slightly earlier than those taxa requiring ice-free terrain. These expansion events contrast with the relatively stable demographic histories inferred for four penguin taxa: *Eudyptes moseleyi*, *E. chrysocome*, *E. pachyrhynchus* and *E. robustus* persisting in ice-free habitats. Shallow genetic structure was also detected in all ice-affected taxa across the vast Southern Ocean, consistent with rapid post-glacial colonisation of sub-Antarctic and Antarctic shores. Together, these analyses demonstrate dramatic, ecosystem-wide responses to Southern Ocean climate change, and highlight potential for rapid biological shifts in Antarctica as climate warming continues.

Introduction

Climate change is substantially impacting the abundance and distribution of the planet's wildlife, with many species' ranges shifting poleward as a result of climate warming (Chen *et al*., 2011). Large-scale shifts also occurred following the end of the LGM (18 – 25 kya) (Hewitt, 2000; Davis & Shaw, 2011; Reid *et al*., 2019), with temperate refugial populations of many

species expanding into the high latitudes. While such range shifts are apparently readily achieved in Northern Hemisphere continental regions where terrestrial habitats are more continuous (Parmesan & Yohe, 2003), the challenges of climate change can be particularly pronounced for isolated and fragmented populations which rely on long-distance dispersal (Trakhenbrot *et al*., 2005). Importantly, many high-latitude coastal and terrestrial ecosystems of the Southern Hemisphere are isolated by substantial oceanographic barriers (*Figure 14*; see Clucas *et al*., 2018 and Frugone *et al*., 2018), with several Southern Ocean circumpolar fronts (such as the STF and the APF), potentially representing physical and thermal barriers to southward range expansion of some species (Fraser *et al*., 2011).

Understanding past shifts in species distributions is crucial for forecasting responses to contemporary and future climate change. Currently, there is considerable uncertainty surrounding the extent to which high-latitude populations might have persisted in the Southern Ocean throughout the LGM versus the extent of post-LGM range expansion (Fraser *et al*., 2009; Fraser *et al*., 2011, González-Wevar *et al*., 2011, González-Wevar *et al*., 2012, González-Wevar *et al*., 2013, González-Wevar *et al*., 2016, Fraser *et al*., 2013; Younger *et al*., 2016; Carrea *et al*., 2019; Chau *et al*., 2019). Recent data hint at major ecosystem-wide change that followed reductions in winter sea-ice (Burg & Croxall, 2001; Ritchie *et al*., 2004; Fraser *et al*., 2009; Clucas *et al*., 2014; Trucchi *et al*., 2014; Younger *et al*., 2015a; Younger *et al*., 2015b; Younger *et al*., 2015c; Cristofari *et al*., 2016; Younger *et al*., 2016; Cristofari *et al*., 2018; Rexer-Huber *et al*, in press). Past biological expansions into habitats freed from receding ice can potentially be reconstructed via genetic analysis of modern populations (Hewitt, 2000; Waters *et al*., 2013; Rexer-Huber *et al*., in press). Crucially, while several studies of Southern Ocean species have detected relatively shallow genetic structure consistent with recent demographic shifts (e.g. Fraser *et al*., 2009, Cristofari *et al*., 2016; Younger *et al*., 2016; Clucas *et al*., 2018; Cristofari *et al*., 2018; Frugone *et al*., 2018; Chau *et al*., 2019; Rexer-Huber *et al*., in press), a comprehensive genome-wide assessment of Southern Hemisphere assemblages is lacking. Moreover, responses to climate change may vary substantially among taxa (Carrea *et al*., 2019; Maggs *et al*., 2008; Stewart *et al*., 2010). Therefore, distinguishing between concerted (multi-species assemblages) shifts versus idiosyncratic (single-species) changes is needed to understand how Southern Hemisphere ecosystems might respond to future anthropogenic climate change (Walther *et al*., 2012).

Penguins are iconic flightless marine birds that inhabit coastlines of all major Southern Hemisphere landmasses, with much of their species diversity currently located in Antarctica and the surrounding sub-Antarctic islands (south of the STF; see *Figure 14* and *Supplementary*

Figure 13). Although most penguins are philopatric to their natal colonies (García Borboroglu & Boersma, 2013), some disperse vast distances, sometimes traversing major oceanographic fronts (see also Mattern *et al*., 2018). As penguins spend much of their lives at sea (Thiébot *et al*., 2012) they represent important components of both coastal and marine ecosystems across the Southern Ocean (Woehler *et al*., 2011). In this study, we test for concerted biological responses to climate change by analysing several thousand SNPs across 12 Antarctic, sub-Antarctic and temperate penguin taxa. We detect strong genomic signatures of population expansions in all penguin taxa currently distributed largely within the LGM sea-ice zone, consistent with concerted re-colonisation of Antarctic and sub-Antarctic coasts during post-LGM climate warming. These southern expansion events are in stark contrast to relatively stable recent demographic histories inferred for four temperate penguin taxa. Our results suggest consistent population dynamics in response to postglacial ice reduction, and demonstrate the potential for rapid change to Southern Ocean ecosystems under future climate warming.

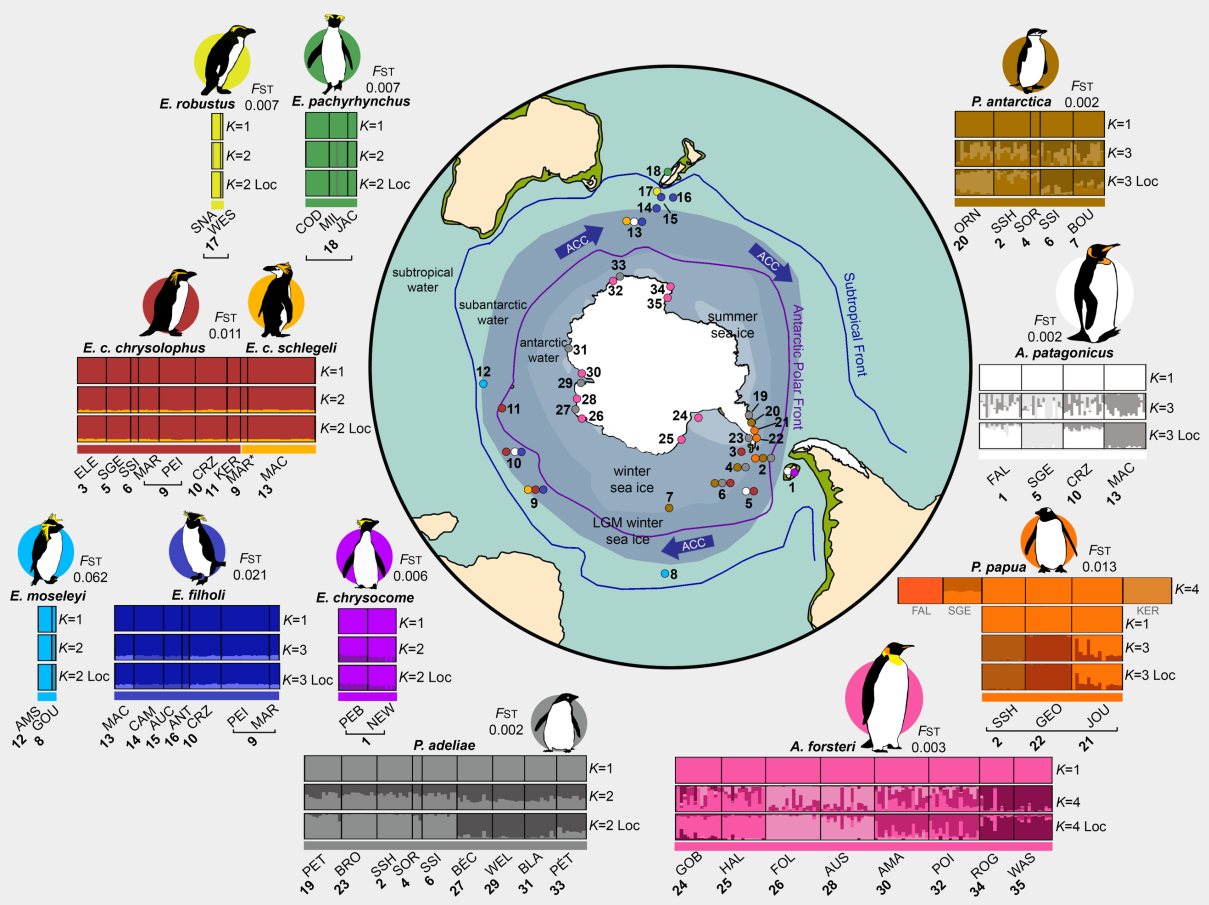


Figure 14. Sampling locations and genetic structure of 12 *Eudyptes*, *Pygoscelis* and *Aptenodytes* penguins. The map (adapted from Fraser *et al*., 2011) shows the ACC, the STF, the APF, modern summer and winter sea-ice, inferred winter sea-ice during the LGM (see Gersonde *et al*., 2005, Fraser *et al*., 2009 and Fraser *et al*., 2011), sea-level during the LGM (green) and glaciation during the LGM (white). Structure plots for all 12 taxa are displayed; the top plot for each taxon represents the most likely number of genetic clusters (*K*), the middle plot for each taxon show a higher value of *K* that may

represent subtle genetic structure (without location priors), and the bottom plot for each taxon represents the same higher value of *K*, with location priors (Loc). The coloured bar underneath the bottom plots indicates which population belongs to a given taxon (e.g. for *Eudyptes chrysolophus chrysolophus* and *E. c. schlegeli*) or which populations were analysed in downstream analyses (e.g. *Pygoscelis papua*). Structure plots for *Aptenodytes* and *Pygoscelis* penguins are adapted from Clucas *et al*. (2016), Younger *et al*. (2017) and Clucas *et al*. (2018). With the exception of *P. papua*, all Structure analyses demonstrated a most likely *K* of one, with low population differentiation (F_{ST}) values (global F_{ST} is shown beside each taxon). Numerical codes for each sampling location is noted on the map and underneath each structure plot; FAL, PEB and NEW are Falkland Islands; SSH is South Shetland Islands; ELE is Elephant Island; SOR is South Orkney Islands; SGE is South Georgia; SSI is South Sandwich Islands; BOU is Bouvet; GOU is Gough Island and Tristan da Cunha; MAR is Marion Island, PEI is Prince Edward Islands; CRZ is Crozet; KER is Kerguelen; AMS is Amsterdam island; MAC is Macquarie Island; CAM is Campbell Island; AUC is Auckland Islands; ANT is Antipodes Islands; SNA is The Snares, WES is Western Chain; COD is Codfish Island, MIL is Milford Sound, JAC is Jackson Head; PET is Peterman Island; ORN is Orne Harbour; JOU is Jougla Point; GEO is George's Point; BRO is Brown Bluff, GOB is Gould Bay; HAL is Halley Bay; FOL is Fold Island; BÉC is Béchervaise Island; AUS is Auster; WEL is Welch Island; AMA is Amanda Bay; BLA is Blakeney Point; POI is Point Géologie, PÉT is Pétrels Island; ROG is Cape Roget and WAS is Cape Washington. Refer to *Supplementary Figure 13* for specific sampling locations.

Materials and Methods

We generated a novel *Eudyptes* dataset. Following filtering (see below) this dataset spans *E. chrysolophus chrysolophus*/*E. c. schlegeli* (n = 125), *E. moseleyi* (n = 8), *E. chrysocome* (n = 26), *E. filholi* (n = 71), *E. pachyrhynchus* (n = 21) and *E. robustus* (n = 5), comprising thousands of SNPs that were generated using Diversity Arrays Technology Pty Ltd (DArT-seq™) technology (see *Supplementary Table 14*). In addition to our novel *Eudyptes* dataset, we also obtained filtered Restricted site Associated DNA Sequencing (RAD-seq) datasets from an additional five penguin taxa that were generated by Clucas *et al*. (2016), Younger *et al*. (2017) and Clucas *et al*. (2018), comprising *Pygoscelis antarctica* (n = 44), *P. adeliae* (n = 87), *P. papua* (n = 67), *Aptenodytes forsteri* (n = 110) and *A. patagonicus* (n = 64) (see *Supplementary Tables 15 – 16*).

Sampling and DNA Extraction

Blood, tissue or feathers were obtained from 428 wild *Eudyptes* penguins encompassing the breeding distributions of seven taxa (see above; *Supplementary Figure 13* and *Supplementary Table 14*). DNA was extracted from each sample using the Qiagen DNeasy® Tissue Extraction Kit (Qiagen Inc., Chatsworth, California, USA) at MWLR, Lincoln or at the University of Oxford, Oxford. Several modifications were made to the user protocol, depending on the tissue type (blood suspended in QL buffer, blood suspended in EtOH, freeze-dried blood cells, tissue or feathers). Specifically, we (1) increased the amount of ProtK to 40 μL and added 5 μL of RNase to all samples; (2) we did not include any PBS to samples that had been suspended in QL buffer; (3) for freeze-dried samples, we increased the lysis incubation to 45 min and

increased the purification incubation to 15 min; (4) for feathers, we increased the digestion to 24 h at 56°C, and then raised the temperature to 60°C for the final 15 min of the digestion; (5) after adding 'AL' buffer (Qiagen), we incubated all samples at 70°C for 45 min, and precipitated the DNA in cold 100% EtOH, and (6) we eluted the feather samples in 100 µL of 'AE' (Qiagen) buffer twice; first at 70°C for 15 min, and then we recycled the buffer to the spin column membrane, incubating for an additional 5 min. We did not modify the user protocol for tissue samples. The quality and quantity of each DNA sample was assessed using a qubit fluorometer, or on a 1% agarose gel in 0.5 X TBE. We diluted samples to a final concentration of 50 – 100 ng/µL. For those samples that were already too dilute, we combined multiple DNA extractions from the same individual (see *Supplementary Table 14*), and concentrated the combined DNA using a speedvac to reduce the volume to a minimum of 10 µL. We tested for buffer-activated nucleases in our DNA extractions by combining 1 µL of each DNA sample with 10 X restriction enzyme buffer and 8 µL RO H_2O, incubating each sample at 37°C for 2 h. Following incubation, the quality of each sample was visualised on a 0.8% agarose gel in 0.5 X TBE. 282 of the highest quality samples, which encompassed all *Eudyptes* taxa (except *E. sclateri*), were retained for library preparation.

DArT-Seq™ library preparation and SNP discovery

Library preparation for SNP discovery was characterised using DArT-seq™ in Canberra, Australia. DArT-seq™ represents a combination of DArT complexity reduction methods and NGS platforms (Kilian *et al*., 2012; Courtois *et al*., 2013; Cruz *et al*., 2013; Raman *et al*., 2014). DArT-seq™ is a relatively recent addition to established genomic sequencing methods, such as RAD-seq (Baird *et al*., 2008) and Genotyping by Sequencing (GBS; Elshire *et al*., 2011), and has been used in several population genomic studies of wild organisms (e.g. Morales *et al*., 2018; de Oliveira *et al*., 2019). DArT-seq™ SNP selection is optimised for each organism by selecting the most appropriate complexity reduction method, including the size of the representation and the fraction of the genome selected for assays. For our *Eudyptes* samples, the Pstl-Hpall method was chosen. Each DNA sample was processed following Kilian *et al*. (2012), but replacing a single Pstl-compatible adaptor with two additional adaptors that corresponded to two different restriction enzyme overhangs (Sansaloni *et al*., 2011). The Pstl compatible adaptor was designed to include the Illumina flow-cell attachment sequence, the sequencing primer sequence and the barcode region (similar to GBS; see Elshire *et al*., 2011). The reverse adaptor contained a flow-cell attachment region and the Hpall-compatible overhanging sequence. Initial denaturation during PCR was undertaken at 94°C for 1 min, followed by 30 cycles of 94°C for 20 s, 58°C for 30 s, 72°C for 45 s, with a final extension at

72°C for 7 min. Following PCR, equimolar amounts of amplification products from each sample were bulked and applied to a c-Bot (Illumina) bridge PCR, which was run for 77 cycles, and was followed by single-read sequencing across three lanes on an Illumina Hiseq 2500.

Sequences generated from each lane were initially processed using an in-house DArT analytical pipeline. Poor quality sequences were removed from the fastq files, while applying stringent criteria to the barcode region, in comparison to the rest of the genome. This ensured that the assignments of the sequences to our specific DNA samples were reliable. Approximately 1,290,000 (plus or minus seven percent) sequences per library, and 1,550,000 sequences per sample were used in the marker calling. To ensure repeatability, the bioinformatic procedure for 53 samples was replicated. Specific filtering included; (1) filtering of the barcode region, with a minimum Phred mass score of 30, and a minimum pass percentage of 75; (2) filtering of the whole read quality, with a minimum Phred score of 10, and a minimum pass percentage of 50, while identical sequences were collapsed into the fastqcall file. The fastqcall file was then used in a secondary in-house pipeline for DArT proprietary SNP and SilicoDArT calling algorithms, which included the presence or absence of restriction fragments in representation. All unique sequences from the set of Fastqcol files were clustered by sequence similarity at a distance threshold of 3 bp. The sequence clusters were then parsed into SNP and silicoDArT markers, while utilising a range of metadata parameters derived from the quantity and distribution of each sequence across all samples, combined with previous experience of Mendelian behaviours of DArT-seq™ markers. Given a high level of technical replication was included in the DArT-seq™ genotyping process, reproducibility scores were able to be calculated for each candidate marker. The candidate markers output by ds14 were further filtered on the basis of the reproducibility values (>95%), average count for each sequence (based on the sequencing depth) and the call rate (the proportion of samples for which the marker was scored). Only 21 samples failed the sequencing pipeline, so our resulting dataset consisted of 134,907 SNPs across 261 *Eudyptes* penguins (see *Supplementary Table 14*).

To increase the sample size for each *Eudyptes* colony, we grouped samples that were collected from the same island archipelago (see *Supplementary Table 14*). We used the DartR vignette v1.1.6 (Gruber *et al*., 2018) in R v.3.5.1 (R Core Team, 2018) to prepare the DArT-seq™ data for downstream analyses. There are occasional taxonomic uncertainties between closely related *Eudyptes* taxa (Banks *et al*., 2006; Jouventin *et al*., 2006; Christidis & Boles, 2008; Jouventin & Dobson, 2017; Frugone *et al*., 2018; Frugone *et al*., 2019; Mays *et al*., in press; see also Cole *et al*., 2019b; *Chapter 2*; Cole *et al*., 2019a; *Chapter 4*; Cole *et al*., 2018; *Chapter 5*). Therefore we filtered 10 datasets separately: (1) all samples; (2) *E. chrysolophus chrysolophus* and *E. c.*

schlegeli; (3) *E. filholi*, *E. chrysocome* and *E. moseleyi*; (4) *E. filholi* and *E. chrysocome*; (5) *E. filholi*; (6) *E. chrysocome*; (7) *E. moseleyi*; (8) *E. pachyrhynchus* and *E. robustus*; (9) *E. pachyrhynchus*; and (10) *E. robustus* (see *Supplementary Table 15*). Specifically, we filtered (1) reproducibility (gl.filter.repavg; t=1, a measure applicable only to DArT-seq™); (2) removed invariant sites that had resulted from the removal of populations or individuals when creating each initial dataset (*gl.filter.monomorphs*); (3) all loci that had call rates <95% (*gl.filter.callrate*; method '*loc*'); (4) discarded all individuals with a call rate of <90% (*gl.filter.callrate*; method '*ind*'); (5) all loci with trimmed sequence tags that could be paralogues (*gl.filter.hamming*); and (6) all loci that exhibited departures from Hardy Weinberg Equilibrium in any colony (*gl.filter.hwe*; $P = 0.05$ following Bonferroni correction). DartR, PGDSpider (Lischer & Excoffier, 2012) and easySFS (https://github.com/isaacovercast/easySFS) were used to convert our neutral set of SNPs to other formats for downstream analyses. Analyses were undertaken on the NZ eScience Infrastructure (NeSI) and the CIPRES scientific gateway.

Phylogenomic analysis of single nucleotide polymorphism data

The use of SNP datasets in phylogenetics has been recently assessed, with overall support for the method (Gonen *et al*., 2015; DaCosta & Sorenson, 2016), especially for studies involving recently diverged taxa, such as *Eudyptes*. RAxML remains one of the best methods for inferring ML phylogenies (Stamatakis, 2006; Stamatakis, 2014). To clarify the evolutionary relationships among our *Eudyptes* penguin samples, we created a ML phylogeny using RAxML-HPC v.8.2.1 (Stamatakis, 2006; Stamatakis, 2014). We followed phylogenetic methods in Buckley *et al*. (2018). Specifically, we removed all missing sites, and all heterozygous states were resolved by randomly assigning one or the other SNP variant to the individual using the DartR function *gl2fasta* in R. We then used Phrynomics v.2.0 (https://github.com/bbanbury/phrynomics) to ensure that all invariable sites had been removed, resulting in a final alignment consisting of 2241 SNPs. We created 20 independent ML tree inferences, drawing bootstrap support from 1000 replicates on the best scoring topology. Searches were conducted under the GTRGAMMA nucleotide substitution model, and we applied an ascertainment bias using the Lewis correction model. The best scoring ML tree was visualised using FigTree v.1.4.2 (http://tree.bio.ed.ac.uk/).

Population genomic summary statistics

We undertook similar population summary statistics and structure analyses for *Eudyptes* as has previously been implemented for *Pygoscelis* and *Aptenodytes*, based on the RAD-seq data

(Clucas *et al*., 2016; Younger *et al*., 2017; Clucas *et al*., 2018). Summary statistics were calculated from the separate datasets containing each *Eudyptes* taxon (*Supplementary Table 15* and *17*). Within each colony, the number of private alleles were calculated using poppr v.2.8.1 (Kamvar *et al*., 2014; Kamvar *et al*., 2015) in R. The observed (H_O) and expected heterozygosity (H_E) and the inbreeding coefficient (G_{IS}) were calculated using Genodive v2.0b27 (Nei, 1987; Meirmans *et al*., 2014). In addition, we calculated AMOVA-based global population differentiation (F_{ST}) indices (Excoffier, 1995) for each of the 10 datasets using 999 permutations in Genodive. Finally, we used the StAMPP v.1.5.1 package (Pembleton *et al*., 2013) in R to measure pairwise F_{ST} for each dataset. 95% confidence intervals (CIs) and *P*-values were generated by calculating F_{ST} over 10,000 bootstraps. To account for multiple testing, we corrected the critical *P*-values using the Holm-Bonferroni sequential criterion (Holm, 1979).

Population genomic structure

Genetic clusters within each *Eudyptes* dataset were visualised using three methods. We first implemented a principal coordinates analysis (PCoA) using adegenet (Jombart & Collins, 2015). We then used Structure v.2.3.4 (Pritchard *et al*., 2000) via Structure Threader (https://github.com/StuntsPT/Structure_threader; Pina-Martins *et al*., 2017) to model population structure under a Bayesian framework using the admixture model with allele frequencies correlated across populations (e.g. the 'F' model; Falush *et al*., 2013). Each analysis was conducted 20 times for 150 million MCMC steps, with the first 50,000 iterations discarded as burnin. With the exception of the global dataset that encompassed all species, each analysis was run twice, with and without location priors. We used the Evanno method (Evanno *et al*., 2005) to estimate the most likely number of genetic clusters (*K*) for each dataset (*Supplementary Table 18*), running K values from 1 to $n + 2$ (where n is the number of populations sampled per species), using Structure Harvester core version vA.2 (Earl & von Holdt, 2012). Exploring different values of *K* can be helpful to gain insights into different levels of genetic structure (e.g. Gilbert *et al*., 2012; Janes *et al*., 2017), especially in populations with limited structure (e.g. *Aptenodytes forsteri*: Cristofari *et al*., 2016, Younger *et al*., 2017 and *Eudyptes chrysolophus chrysolophus*/*E. c. schlegeli*: Frugone *et al*., 2018, Frugone *et al*., 2019). Therefore, we present Structure results that represent different values of *K*. In addition, we undertook a discriminant analysis of principal components (DAPC) of each dataset using adegenet. We ran the DAPC analysis a second time for each dataset, with individuals grouped by their colony of origin. Cross-validation with 1000 replicates was used to determine the most appropriate number of PCs to retain, as suggested by Jombart and Collins (2015).

Introgression between species and colonies

We tested for introgression between *Eudyptes* colonies, and within *Pygoscelis adeliae* and *P. papua* by employing the SNAPP tree set analyser implemented in BEAST v.2.4.7 (Bouckaert *et al*., 2014). We limit our analysis to the above taxa as Clucas *et al*. (2016), Younger *et al*. (2017) and Clucas *et al*. (2018) previously undertook SNAPP analyses for *Aptenodytes patagonicus*, *A. forsteri* and *Pygoscelis papua* (see *Supplementary Figure 14*). For *Eudyptes*, we grouped the most closely related taxa together: (1) *E. c. chrysolophus*/*E. c. schlegeli*; (2) *E. moseleyi*, *E. chrysocome* and *E. filholi*; and (3) *E. pachyrhynchus* and *E. robustus*, based on our Structure results and the taxonomic literature (as above). Mutation rates were calculated based on the data in the alignment. We ran the analysis up to 100 million MCMC samples, logging parameters every 10,000 trees with a pre-burnin of 5000. We assessed convergence using Tracer v.1.6.0 (http://beast.bio.ed.ac.uk/) until ESS values were greater than 200. Because SNAPP is computationally demanding, each *Eudyptes* analysis was conducted three times and each *Pygoscelis adeliae* and *P. antarctica* analysis was run twice, while assigning two different individuals per colony to each of the replicates (*Supplementary Table 19*). As such, we were better able to detect if there are fine-scale differences (e.g. genetic structure) that may influence the topology. DensiTree v.2.4.7 (Bouckaert & Heled, 2014) was used to visualize the posterior distributions of topologies as cloudograms.

Testing for demographic expansions in *Eudyptes*, *Pygoscelis* and *Aptenodytes*

We reconstructed the demographic histories for the 12 penguin taxa over the 1,000,000 years, by estimating the time and magnitude of demographic changes using five different approaches (*Eudyptes moseleyi* and *E. robustus* were excluded from some analyses due to low sample size). Specifically, we reconstructed the demographic histories using CubSFS (Waltoft & Hobolth, 2018); used Fastsimcoal2 (Excoffier *et al*., 2013) to estimate the effective population size (N_e) at five different late-Quaternary time periods; obtained Tajima's D; identified the change in theta values as inferred by our previous SNAPP analyses; and tested for synchronous expansion using Multi-dice (Xue & Hickerson, 2017).

Previous studies have suggested that there may be shallow genetic structure within *Pygoscelis* and *Aptenodytes* taxa (Clucas *et al*., 2014; Younger *et al*., 2015a; Clucas *et al*., 2016; Cristofari *et al*., 2016; Levy *et al*., 2016; Vianna *et al*., 2017; Younger *et al*., 2017; Clucas *et al*., 2018; Cristofari *et al*., 2018) (see *Figure 14*). We interpret this shallow genetic structure (based on low F_{ST} and estimated K of 1) to represent a scenario of high gene flow, and consider *Pygoscelis antarctica*, *P. adeliae*, *Aptenodytes forsteri* and *A. patagonicus* as single panmictic populations.

To account for the genetic structure observed in *P. papua* (e.g. four distinct lineages; Clucas *et al*., 2018; see *Figure 14*), we focussed our CubSFS analysis on individuals from the South Shetland Islands and the West Antarctic Peninsula (Jougla Point and George's Point) (see *Figure 14*). Filtering of these *P. papua* individuals was undertaken using VCFtools (Danecek *et al*., 2011) and applying a minor allele frequency (MAF) of 0.01.

We applied further filtering to ensure our datasets (DArT-Seq versus RAD-Seq) were consistent with one another (except SNAPP, see below). We used VCFtools (Danecek *et al*., 2011) to apply further filtering to five DArT-seq *Eudyptes* datasets (*E. chrysolophus chrysolophus*/*E. c. schlegeli*, *E. chrysocome*, *E. moseleyi*, *E. filholi*, *E. pachyrhynchus*; *E. robustus* was not included due to low sample sizes), and the RAD-seq datasets. Specifically, we filtered the *Eudyptes* datasets with a MAF of 0.01, reflecting the filtering described in Clucas *et al*. (2018). For the *Aptenodytes* and *Pygoscelis* datasets, we filtered loci with call rates <95%, and individuals with call rates <90%, reflecting the filtering for the *Eudyptes* samples (consistent with the data obtained in this study). For all demographic analyses, we considered *E. c. chrysolophus* and *E. c. schlegeli* as conspecific, based on our genetic structure and phylogenomic results.

Testing for demographic expansions using CubSFS

We used CubSFS v1.0 to test for demographic expansions among all taxa, except *Eudyptes robustus* due to limited sample sizes. For each *Eudyptes, Pygoscelis,* and *Aptenodytes* dataset, we projected the folded allele frequency spectrum using EasySFS (https://github.com/isaacovercast/easySFS). We then adjusted the number of invariant sites in our allele frequency spectrum to reflect the total number of invariant loci characterized within each species. We read the folded, MAF filtered (0.01) SFS into R v3.5.1 using tidyverse v1.2.1 (https://github.com/tidyverse/tidyverse.org) functions. After obtaining the sample size and total number of sites from the SFS, the invariant site category was removed, and CubSFS v1.0 was used to estimate the demographic history for each taxon, using 29 knots and a t_m of 0.25 coalescent units. Because of variance in how the invariant site category was recorded per dataset due to the RAD-seq methodologies employed in this study (see Clucas *et al*. 2018), we generated the expected SFS and a negative exponential model was used to estimate the 'true' number of invariant sites. We also amended the observed SFS site categories likely affected by the MAF by extracting the expected SNP counts for these sites. We then used CubSFS to re-estimate the demographic history from the amended SFS (as above). We then generated 10 bootstrap replicates per taxa and used CubSFS to estimate the demography for each of these replicates with the same parameters as used for the amended SFS.

We then obtained demographic reconstructions in units of N_e against year by using the estimated number of invariant sites, the per-generation mutation rate, and species-specific generation times. We assumed an average generation time of 8 – 14 years depending on the taxon (based on Forcada & Trathan, 2009). For *Eudyptes chrysolophus chrysolophus* and *E. c. schlegeli*, we averaged the two generation times presented in Forcada and Trathan (2009). For the taxa that did not have a generation time in Forcada and Trathan (2009), we used the same generation time as the most closely related taxon, which was always in the same genus. As such, we used a generation time of eight years for all taxa except *Aptenodytes patagonicus* which has a generation time of nine years, and *A. forsteri* which has a generation time of 14 years. We used a mutation rate of 2.6×10^{-7} per site per generation (Trucchi *et al*., 2014; Cristofari *et al*., 2016). These demographic histories were plotted using R, and the tidyverse, gridExtra (https://cran.r-project.org/web/packages/gridExtra/index.html) and scales (https://cran.r-project.org/web/packages/scales/index.html) libraries.

Testing for demographic expansions using Fastsimcoal2

We used Fastsimcoal 2.5 to estimate fluctuations in N_e over the past 1 million years. For each taxon, we estimated N_e at five time points fixed at 1 mya, 140 kya, 50 kya, 15 kya and 1 kya BP, which correspond to mid-Pleistocene, the Penultimate glaciation, the last glaciation, the end of the LGM and the late-Holocene. We chose those time points based on both the major climatic shifts (Lisiecki & Raymo, 2005) and the long-term historical demographic trajectories from the CubSFS analysis. We used uniform priors for N_e (1×10^3 - 2×10^6) and assumed a generation time of 8 – 14 years (as above).

We first performed 50 replicate runs in Fastsimcoal2 using the following parameters: -n 100000 -m -M 0.01 -L 40. Similar to the CubSFS analysis, the monomorphic site category was removed using the --removeZeroSFS option when performing these 50 replicates. We then retained the set of parameters with the highest composite likelihood as the best point estimate. In order to account for variance in how the monomorphic site category was recorded per dataset due to the differing RADseq methodologies, we then used the expected SFS generated for the model that obtained the highest composite likelihood and used a negative exponential model to estimate the 'true' number of monomorphic sites. We also amended the observed SFS site categories likely affected by the MAF by extracting the expected SNP counts for these sites. Next, we used this corrected SFS and updated the total number of sites in our models to estimate the final likelihood of the preferred model by running 50 replicates of 100,000 simulations with the best point estimate. Finally, to estimate the 95% confidence intervals of each N_e we used a parametric bootstrap procedure by creating 100 bootstrap runs simulated from the SFS of the

point estimates. The data-type was changed to DNA (1 bp), with the number of chromosomal segments equalling the total number of loci in the SFS (including monomorphic sites). We ran 50 replicates per bootstrap for a total of 5000 data points to estimate posterior distributions of N_e. As our *Eudyptes* SNP dataset was generated using a different protocol to the *Aptenodytes* and *Pygoscelis* SNP datasets, we avoided potential bias in our results by calculating the change in N_e based on the percentage of N_e at time 1 Ma.

Testing for demographic expansions using Tajima's D

Tajima's D was computed using δaδi (Gutenkunst *et al*., 2009) from the downprojected allelic frequency spectrum, ignoring allele frequencies below a MAF of 0.01. Confidence intervals were obtained by re-sampling 1000 SNPs in each dataset using in-house scripts.

Testing for demographic expansions using SNAPP

We also investigated the change in N_e in each of our taxa as represented by the change in theta between the terminal lineages and nearest internal branches derived from the previous SNAPP analyses. For the *Eudyptes* taxa, *Pygoscelis adeliae* and *P. antarctica*, we used results from the SNAPP analyses described above. For the remaining taxa, we used previously published SNAPP analyses obtained from Clucas *et al*. (2016) for *Aptenodytes patagonicus*, Younger *et al*. (2017) for *A. forsteri* and Clucas *et al*. (2018) for *Pygoscelis papua*. While CubSFS, Fastsimcoal2 and Tajima's D explored population expansions across the combined sample locations for each taxon, we explored the change in theta inferred from SNAPP separately for each sampling locality (e.g. for *P. papua*, the four lineages represented in *Figure 14*; see also Clucas *et al*., 2018). To do this, we constructed maximum-clade credibility trees for each taxon using common ancestor heights through TreeAnnotator v2.5.0. This maximum-clade credibility tree was used as input into the SNAPP tree set analyser with 10% burnin to calculate the theta values for each terminal tip (e.g. the sampling location) and the nearest internal ancestral branches. Using the maximum-clade credibility trees, we also estimated the length of time (in years) that the change of theta had occurred across for each sampling location by extracting the age of the nearest ancestral node (in number of substitutions), dividing this by the per-generation mutation rate (see above) and then multiplying by the estimated generation time (in years) for each taxon (see above).

Testing for demographic expansions using Multi-dice

Finally we used the R package Multi-dice (Xue & Hickerson, 2017) to test for a population expansion among the expanding penguin taxa following the LGM. We tested scenarios of one

co-expansion within the last 50 Ka that could be synchronous or not across the 'expanding' taxa (*Eudyptes chrysolophus chrysolophus*/*E. c. schlegeli*, *E. filholi*, *Pygoscelis adeliae*, *P. papua*, *P. antarctica*, *Aptenodytes patagonicus* and *A. forsteri*, based on the combined results of CubSFS, Fastsimcoal2, Tajima's D and SNAPP). We did not analyse *Eudyptes moseleyi*, *E. pachyrhynchus*, *E. chrysocome* or *E. robustus* as we consider those taxa to have had a relatively stable population size since the LGM (based on the combined results of CubSFS, Fastsimcoal2, Tajima's D and SNAPP). We projected all SNPs datasets to 40 haploid samples using easySFS, maximising the variation in *E. pachyrhynchus* which had the smallest number of sampled individuals and equalising the SFS across species as required by Multi-dice. We performed 100,000 simulations under a simple scenario comparable to Xue and Hickerson (2017). We allowed for one synchronous expansion event where one to seven co-taxa expand while the others expand idiosyncratically. Expansions were constrained to have a post-expansion N_e within 10 – 100 times ancestral N_e, and for species-specific expansions to occur within 1000 years of each other to be considered as co-expanding. Both the proportions of co-expanding taxa and the timing of the synchronous event were recorded along with the aggregate frequency spectrum (Xue & Hickerson, 2017) for inference. We then sampled the best 5% (5000 simulations) using hierarchical Approximate Cayesian Computation (ABC) as implemented in ABC (Csilléry *et al*., 2012) and the aggregate site-frequency spectrum of each simulation as a single pseudo-observed dataset.

Results

Genetic diversity, structure and admixture of *Eudyptes* penguins

We tested for population genetic structure and introgression (as inferred by Napier, 1968; Woehler & Gilbert, 1990; White & Clausen, 2002; Morrison & Sagar, 2014) between all our *Eudyptes* samples using F_{ST}, RAxML, Structure, DAPC and PCoA (see *Figures 15* and *16* and *Supplementary Table 20*).

As demonstrated by several recent studies (e.g. Frugone *et al*., 2018; Cole *et al*., 2019b; *Chapter 2*; Frugone *et al*., 2019), *E. c. chrysolophus* and *E. c. schlegeli* consistently clustered together with little evidence to distinguish between the two subspecies (*Figure 15* and *16*, bootstrap support: 100; *Supplementary Figure 15*, *Supplementary Tables 20* and *21*), despite clear phenotypic differences (Shaughnessy, 1975; Warham, 1975): *E. c. chrysolophus* are smaller than *E. c. schlegeli* and have a black face, while *E. c. schlegeli* have a white face. Our results consistently support de Dinechin *et al*. (2009), Frugone *et al*. (2018) and Cole *et al*. (2019b; *Chapter 2*) which recognise *E. moseleyi*, *E. chrysocome* and *E. filholi* as three distinct species

(*Figure 15* and *16*, *Supplementary Figures 15* and *16*, *Supplementary Tables 22* – 26), even though *E. chrysocome* and *E. filholi* are still considered subspecies or incipient species (e.g. Mays *et al*., in press). The closely related *E. pachyrhynchus* and *E. robustus* remained difficult to distinguish under our initial global Structure analyses (including $K = 5 – 7$; *Figure 16*), initial DAPC analyses (not shown) and to a lesser extent with RAxML (*Figure 15*; bootstrap support: 92). However, downstream analyses of *E. pachyrhynchus* and *E. robustus* supports recognition of two distinct species (*Figure 16H*, *Supplementary Figures 15* and *16*, *Supplementary Table 27*).

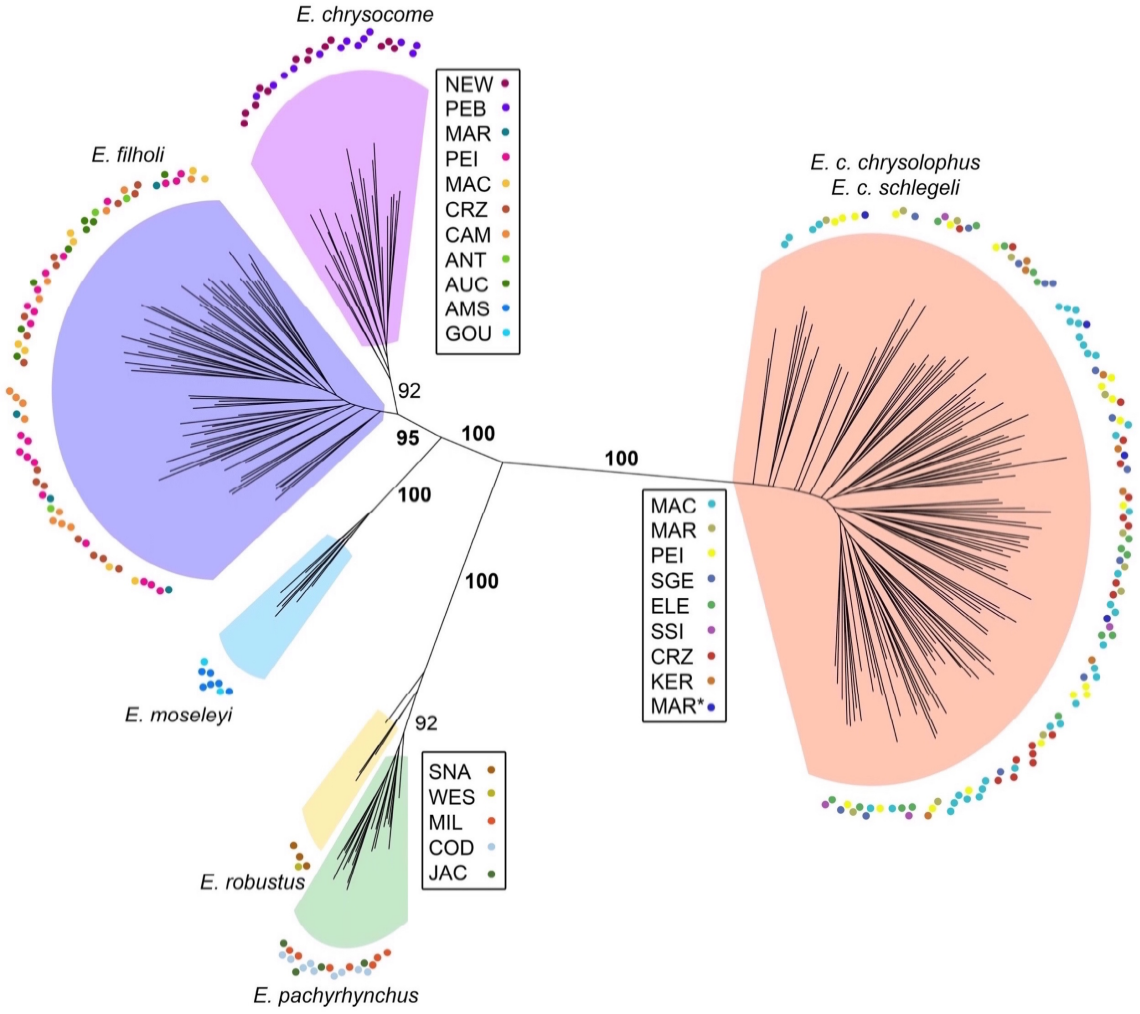


Figure 15. Unrooted ML phylogenetic tree of *Eudyptes* penguins based on SNPs. Bootstrap values for all major clades are shown. Small coloured circles represent the population that individual was derived from. MAR* indicates Marion Island white-faced penguins of *Eudyptes chrysolophus chrysolophus* or *Eudyptes c. schlegeli*. PEB and NEW are Falkland Islands; SGE is South Georgia; SSI is South Sandwich Islands; GOU is Gough Island and Tristan da Cunha; MAR is Marion Island, PEI is Prince Edward Islands; CRZ is Crozet; KER is Kerguelen; AMS is Amsterdam island; MAC is Macquarie Island; CAM is Campbell Island; AUC is Auckland Islands; ANT is Antipodes Islands; SNA is The Snares, WES is Western Chain; COD is Codfish Island, MIL is Milford Sound and JAC is Jackson Head.

With the exception of *E. c. chrysolophus* and *E. c. schlegeli* (which are probably in their earliest stages of speciation; see Cole *et al*., 2019b; *Chapter 2*), our Structure analysis (*Figure 16*)

revealed limited evidence for introgression between the seven *Eudyptes* taxa. Very low levels of introgression (<0.5%) were detected from *E. robustus* to *E. pachyrhynchus*, *E. moseleyi*, *E. filholi* and *E. chrysocome*, from *E. pachyrhynchus* to *E. moseleyi* and *E. filholi*, from *E. moseleyi* to *E. chrysocome* and *E. c. chrysolophus*/*E. c. schlegeli*, from *E. filholi* to *E. robustus*, from *E. chrysocome* to *E. moseleyi*, and from *E. c. chrysolophus*/*E. c. schlegeli* to *E. filholi, E. chrysocome*, *E. pachyrhynchus* and *E. robustus.* Slightly higher levels of introgression (0.6 – 0.9%) were observed from *E. filholi* to *E. moseleyi* (mean 0.6%), *E. chrysocome* to *E. filholi* (mean 0.7%), and from *E. moseleyi* to *E. robustus* (mean 0.9%), while the highest levels of introgression were detected between the most closely related taxa, such as *E. filholi* to *E. chrysocome* (mean 5.1%), between *E. moseleyi* to *E. filholi* (mean 1.%) and between *E. pachyrhynchus* to *E. robustus* (mean 68%, however see *Figure 16*). To explore the patterns among closely related *Eudyptes* taxa and within each taxon, we undertook further analyses on smaller datasets (see *Supplementary Table 15*).

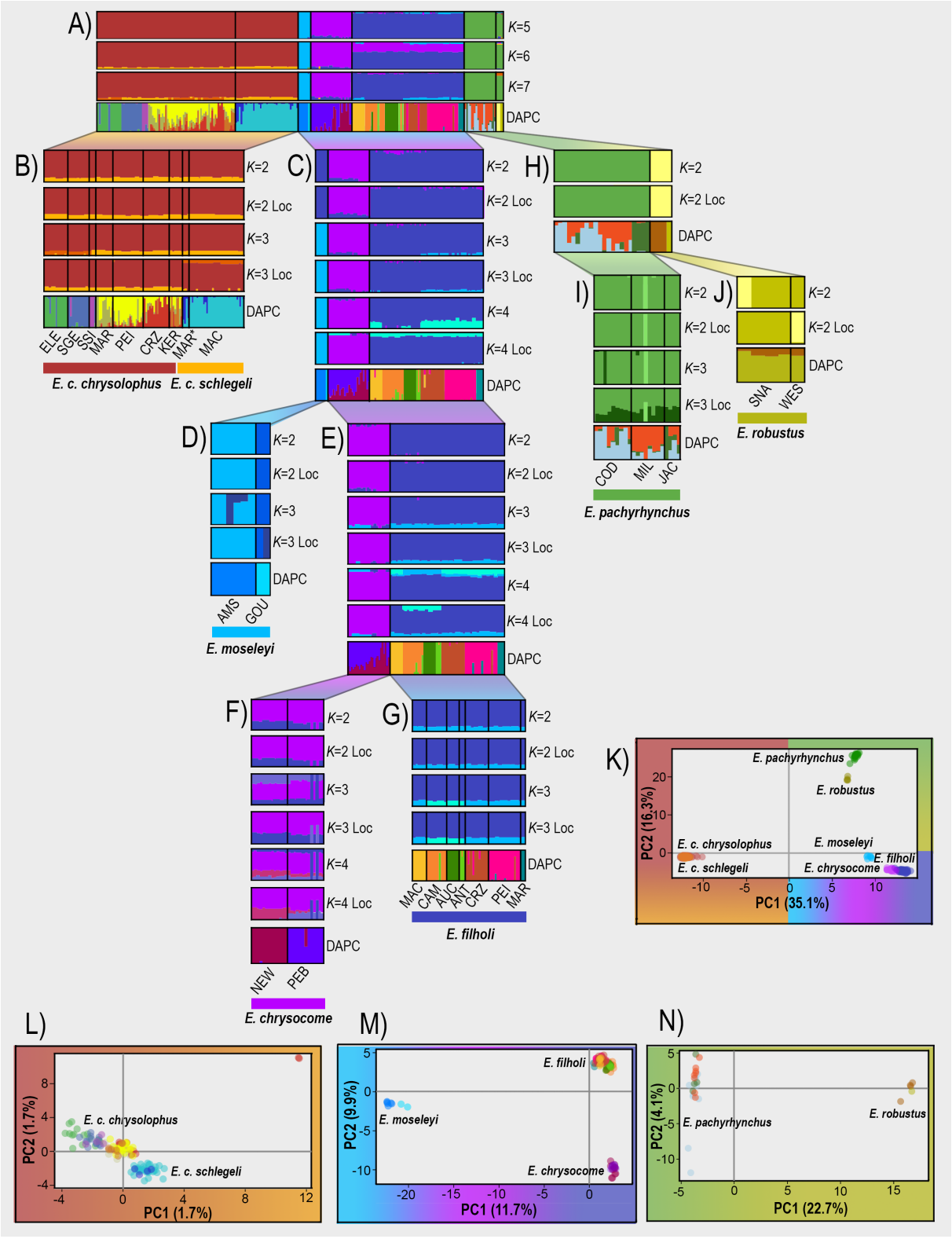


Figure 16. Genetic structure among *Eudyptes* penguins inferred by Structure, DAPC and PCoA. **(A)** – **(J)** (represented in separate blocks of 3 – 7 plots) are Structure and DAPC plots. Each block of plots represents a single dataset, analysed to explore relatedness and introgression within the *Eudyptes* genus (e.g. A), between closely related taxa (e.g. B, C, E and H) and within each taxon (e.g. B, D, F, G, I and J). *K* indicates the number of genetic clusters inferred by that corresponding Structure plot, Loc indicates that the corresponding Structure plot used Location Priors, and DAPC indicates that the corresponding plot was created with DAPC, also using location priors. **(A)** is all *Eudyptes* penguins together (without location priors) for *K* = 5 – 7 (*K* = 1 – 4 and *K* = 8 – 13 are not shown); **(B)** is *E. chrysolophus chrysolophus* and *E. c. schlegeli* for *K* = 2 – 3 (*K* = 1 and *K* = 4 – 11 are not shown, the most

likely *K* is *K* = 1); **(C)** is *E. moseleyi*, *E. chrysocome* and *E. filholi* for *K* = 2 – 4 (*K* = 1 and *K* = 5 – 13 are not shown, the most likely *K* is *K* = 3); **(D)** is *E. moseleyi* for *K* = 2 – 3 (*K* = 1 and *K* = 4 are not shown, the most likely *K* is *K* = 1); **(E)** is *E. chrysocome* and *E. filholi* for *K* = 2 – 4 (*K* = 1 and *K* = 5 – 13 are not shown, the most likely *K* is *K* = 2); **(F)** is *E. chrysocome* for *K* = 2 – 4 (*K* = 1 is not shown, the most likely *K* is *K* = 1); **(G)** is *E. filholi* for *K* = 2 – 3 (*K* = 1 and *K* = 4 – 9 are not shown, the most likely *K* is *K* = 1); **(H)** is *E. robustus* and *E. pachyrhynchus* for *K* = 2 (*K* = 1 and *K* = 3 – 7 are not shown, the most likely *K* is *K* = 2); **(I)** is *E. pachyrhynchus*, for *K* = 2 to *K* = 3 (*K* = 1 and *K* = 4 to *K* = 5 are not shown, the most likely *K* is *K* = 1); and **(J)** is *E. robustus* for *K* = 2 (*K* = 1 and *K* = 3 – 4 are not shown, the most likely *K* is *K* = 1). Each Structure run was repeated 20 times. Structure plots for *K* = 1 are shown in *Figure 14*. The coloured bar underneath the population names indicates which population belongs to which taxon (e.g. for *E. c. chrysolophus* and *E. c. schlegeli*). **(K)** to N) are PCoA; K) encompasses all *Eudyptes* samples and corresponds to the same dataset used to generate the Structure and DAPC plots in A); **(L)** encompasses all *E. c. chrysolophus* and *E. c. schlegeli* samples, and corresponds to the dataset used to generate the Structure and DAPC plots in B); **(M)** encompasses *E. moseleyi*, *E. chrysocome* and *E. filholi* samples, and corresponds to the dataset used to generate the Structure and DAPC plots in C); and **(N)** encompasses the *E. pachyrhynchus* and *E. robustus*, and corresponds to the dataset used to generate the Structure and DAPC plots in H). Eigenvectors one and two are plotted. Each sample in the PCoA corresponds to the colours shown in the DAPC plot. PEB and NEW are Falkland Islands; SGE is South Georgia; SSI is South Sandwich Islands; GOU is Gough Island and Tristan da Cunha; MAR is Marion Island, PEI is Prince Edward Islands; CRZ is Crozet; KER is Kerguelen; AMS is Amsterdam island; MAC is Macquarie Island; CAM is Campbell Island; AUC is Auckland Islands; ANT is Antipodes Islands; SNA is The Snares, WES is Western Chain; COD is Codfish Island, MIL is Milford Sound and JAC is Jackson Head.

Eudyptes chrysolophus chrysolophus and *E. c. schlegeli*

The phylogenetic placement of *E. c. chrysolophus* with *E. c. schlegeli* as a single panmictic population was consistently supported by pairwise F_{ST}, PCoA, Structure, DAPC, and SNAPP (*Figures 15* and *16, Supplementary Figures 15* and *Supplementary Tables 20* and *21*), supporting Cole *et al*. (2018; *Chapter 5*), Frugone *et al*. (2018), Cole *et al*. (2019b; *Chapter 2*), Cole *et al*. (2019a; *Chapter 4*) and Frugone *et al*. (2019). Our Structure analyses suggested the most likely *K* = 2 (*Supplementary Table 18)*. However, the corresponding Structure plot revealed no distinction between *E. c. chrysolophus* and *E. c. schlegeli* (*Figures 14* and *16*), even though the subspecies exhibit clear phenotypic differences. Very subtle population structure between the two taxa was only evident when we examined alternative Structure plots of *K* >3 (*Figure 16*), when we undertook further DAPC analyses using location priors (*Figure 16*) or when we examined PCoA plots (*Figure 16*). Our PCoA and DAPC analysis (when location priors were used; *Figure 16*) revealed the possible presence of three very shallow genetic clusters which fall into three distinct biogeographical regions. These correspond to (1) Elephant, South Georgia and South Sandwich Islands (*E. c. chrysolophus*); (2) Marion, Prince Edward, Crozet and Kerguelen (*E. c. chrysolophus*); and (3) Macquarie and Marion Island 'white-faced penguins' (*E. c. schlegeli*). These shallow clusters were generally supported by our SNAPP analyses (*Supplementary Figure 15*). While our Structure analyses suggest a scenario of high gene flow among these regions, our DAPC analysis (with location priors; *Figure 16*) indicates there may be limited movement of *E. c. chrysolophus* between Elephant,

South Georgia and the South Sandwich Islands, with only few individuals potentially representing migrants (*Figure 16*). Moreover, our results support Frugone *et al*. (2018) who also demonstrated that the 'white-faced' penguins breeding on Marion Island cluster with the Macquarie Island *E. c. schlegeli* individuals (*Figures 14* and *16*), suggesting recent dispersal of *E. c. schlegeli* across the Southern Ocean. As our analyses consistently demonstrated a lack of strong genetic structuring between *E. c. chrysolophus* and *E. c. schlegeli* we consider these subspecies as a single population.

Eudyptes moseleyi, *E. chrysocome* and *E. filholi*

Species-level distinction of *E. chrysocome*, *E. filholi* and *E. moseleyi* was consistent among F_{ST}, PCoA, Structure, DAPC, and SNAPP analyses (*Figures 15* and *16*, *Supplementary Figure 15*, *Supplementary Table 22*), when we examined the three taxa in the same dataset supporting Cole *et al*. (2018 *Chapter 5*), Frugone *et al*. (2018), Cole *et al*. (2019b; *Chapter 2*) and Cole *et al*. (2019a; *Chapter 4*). Our Structure analyses of the three taxa suggested $K = 3$ (*Supplementary Table 18*). However, when we examined the $K = 2$ Structure plot, *E. moseleyi* clustered with *E. filholi* (*Figure 16*). This is surprising, as *E. moseleyi* is basal to *E. filholi*/*E. chrysocome* (see *Figure 9C* and Cole *et al*., 2019b; *Chapter 2*), and because *E. filholi* is occasionally considered a subspecies of *E. chrysocome* (Mays *et al*., in press). This result may reflect the difficulty for Structure to choose two distinct clusters, when three are clearly present. Further Structure results of $K = 3$ (*Figure 16*) revealed possible low levels of introgression within the three taxa. Very low levels of introgression (<0.1%) were observed from *E. chrysocome* to *E. moseleyi* and from *E. filholi* to *E. moseleyi*. Further low levels of introgression were observed from *E. filholi* to *E. chrysocome* (mean 1.9%), from *E. chrysocome* to the *E. filholi* (mean 0.%) and from *E. moseleyi* to *E. filholi* (mean 1.4%). When we examined the $K = 3 – 4$ Structure plots (see *Figure 16*), it appears that only *E. filholi* individuals from the NZ Campbell, Auckland and Antipodes Islands have introgressed with *E. chrysocome*, while all other *E. filholi* samples appear to have experienced introgression with *E. moseleyi.*

Eudyptes filholi

Structure suggested a most likely $K = 1$ for *E. filholi* (*Supplementary Table 18*). However, when we explored the $K = 3$ structure plots (with and without location priors; *Figure 16*), we revealed subtle population structure corresponding to (1) Campbell, Auckland and Antipodes Islands; and (2) Macquarie, Crozet, Prince Edward and Marion Islands (see above and *Figure 16*). This pattern was also reflected in SNAPP, PCoA and to a lesser extent with DAPC (*Figure 16*; *Supplementary Figures 15* and *16*). DAPC (with location priors) revealed limited evidence of movement between *E. filholi* populations; a single individual from Antipodes Island may have

migrated to Campbell Island and there may be some migration between Crozet, Prince Edward and Marion Islands (see *Figure 16*).

Eudyptes chrysocome

Samples for *E. chrysocome* were obtained only from two locations on the Falkland Islands; New Island and Pebble Beach (see *Supplementary Table 14* and *Supplementary Figure 13*). Our Structure analyses suggest a most likely $K = 2$ for *E. chrysocome* (*Supplementary Table 18*). When we examined the $K = 4$ (with location priors) Structure plots, subtle structure between New Island and Pebble Beach was observed (see *Figure 16*). This is also reflected with DAPC (with location priors) (*Figure 16*), and may further indicate that these populations did not experience a population expansion following the LGM (see below). Our PCoA and Structure results revealed two individuals from Pebble Beach as consistent outliers to all other *E. chrysocome* individuals (*Figure 16* and *Supplementary Figure 16*), while DAPC revealed a different individual from Pebble Beach as a possible migrant (*Figure 16*). This may suggest the presence of additional structure between other unsampled *E. chrysocome* colonies.

Eudyptes moseleyi

PCoA, Structure and DAPC suggested there may be population structure between Amsterdam and Gough Islands *E. moseleyi* populations (*Figure 16* and *Supplementary Figure 16*), and may mirror the patterns seen in other taxa breeding north of the LGM sea-ice (e.g. *E. chrysocome*, see above). This pattern could not be analysed in detail due to the low sample size ($n = 6$ for Amsterdam Island and $n = 2$ for Gough Island).

Eudyptes pachyrhynchus and *Eudyptes robustus*

Our initial global Structure and RAxML analyses detected possible introgression between *E. pachyrhynchus* and *E. robustus* (see above; *Figures 15* and *16*). However, all downstream analyses conducted on just the two taxa (*Figure 16*, *Supplementary Figures 15* and *16, Supplementary Table 27*) supported species-level distinctions. This has also been demonstrated by Cole *et al*. (2019a; *Chapter 4*) using COI and CR to analyse the same *E. pachyrhynchus* and *E. robustus* individuals.

Eudyptes pachyrhynchus

We found very little within-species structure for *E. pachyrhynchus* (*Figure 16*, *Supplementary Figure 16*, *Supplementary Table 28*). When we examined the $K = 2 – 3$ Structure plots (with and without location priors; *Figure 16*) a single Milford Sound individual was a consistent outlier. When we examined the $K = 3$ Structure plot (without location priors) an individual from

Codfish Island was as an outlier (*Figure 16*). These individuals are potentially migrants from unsampled colonies.

Eudyptes robustus

Structure, PCoA, DAPC and RAxML found little evidence to suggest population genetic structuring between The Snares and Western Chain *E. robustus* colonies (*Figure 16*, *Supplementary Figure 16*, *Supplementary Table 29*), supporting recent findings by Cole *et al*. (2019a; *Chapter 4*).

Demographic expansions

CubSFS

The demographic histories inferred by CubSFS for 11 penguin taxa (*E. robustus* was excluded due to small sample sizes) revealed concerted demographic expansions in *E. c. chrysolophus*/*E. c. schlegeli*, *E. filholi, Aptenodytes forsteri, A. patagonicus, Pygoscelis papua* and *P. adeliae* (*Supplementary Tables 3 – 30*). All of these recently 'expanded' taxa are predominantly found south of the LGM sea-ice limit (*Figure 14*). Crucially, five of these near-simultaneous demographic expansions began approximately 17.2 thousand years ago (ranging from 20 – 15 thousands years ago) (*Figure 17*, *Supplementary Figures 17* and *18*, *Supplementary Table 30*), at a time of rapid warming in the immediate aftermath of the LGM (Jouzel *et al*., 2007). By contrast, analysis of *Aptenodytes forsteri* shows population expansion earlier than other penguin lineages, suggesting that this ice-adapted species was able to exploit southern regions earlier; see also Li *et al*. (2014) and Cristofari *et al*. (2016). The magnitude of these population expansions is on average a 2.77-fold increase (ranging from 1.18 – 4.40-fold increase; *Figure 17*). In addition, analyses of five taxa (*E. chrysocome*, *E. moseleyi*, *E. pachyrhynchus* and *E. robustus*, and *Pygoscelis antarctica*) predominantly found north of the LGM sea-ice zone (*Figure 14*) revealed decreasing or relatively stable population sizes (*Figures 17 – 18*, *Supplementary Table 30*).

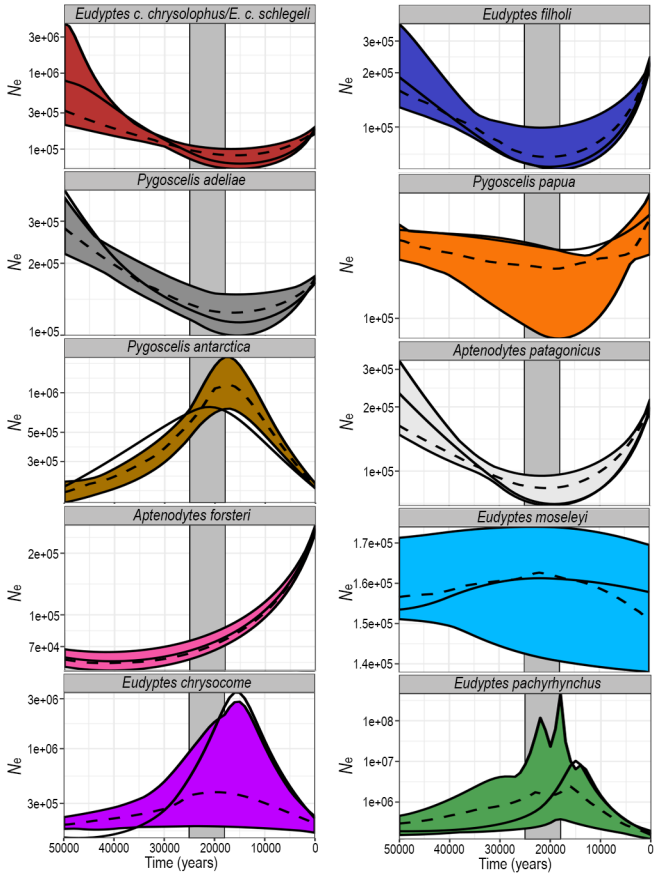


Figure 17. Demographic reconstructions of historical N_e for *Eudyptes*, *Aptenodytes* and *Pygoscelis* penguins inferred by CubSFS. 95% CIs are given by solid colour intervals. Median for bootstrap replicates is given by the dotted line, and the solid line gives the demographic reconstruction for the amended site frequency spectrum. The solid grey bar represents the LGM. Note *Eudyptes robustus* is not included due to small sample size. Note also that *E. chrysolophus chrysolophus* and *E. c. schlegeli* were analysed together.

Fastsimcoal2

Demographic reconstructions of N_e using Fastsimcoal2 (*E. moseleyi* and *E. robustus* were excluded due to low sample sizes) were generally consistent with CubSFS results (*Figure 18*, *Supplementary Figure 19*; *Supplementary Tables 31* and *32*). Based on the mean N_e, all penguin taxa were inferred to have experienced a population decline (range 1.33-fold – 3.07-fold; average 2.07-fold) between 1 mya BP and the Penultimate glaciation (140 kya BP). From the Penultimate glaciation to the Last Glaciation (50 kya BP) all taxa experienced a population expansion (range 2.54 – 8.54-fold; average 5.29-fold). From the Last Glaciation to the LGM (15 kya BP) eight taxa (*Eudyptes chrysolophus chrysolophus*/*E. c. schlegeli*, *E. filholi*, *Pygoscelis adeliae*, *P. papua*, *P. antarctica*, *Aptenodytes patagonicus* and *A. forsteri*) declined (range 1.25 – 9.61-fold; average 4.27-fold), while two taxa (*Eudyptes chrysocome* and *E. pachyrhynchus*) increased (range 1.01-fold – 1.11-fold; average 1.06-fold). From the LGM to the Holocene (1 kya BP) six taxa (*Pygoscelis adeliae*, *P. papua*, *P. antarctica*, *Aptenodytes patagonicus*, *A. forsteri* and *Eudyptes pachyrhynchus*) increased (range 1.02-fold – 12.25-fold;

average 5.79-fold), while four taxa (*Eudyptes chrysolophus chrysolophus*/*E. c. schlegeli*, *E. filholi* and *E. chrysocome*) declined (range 1.06-fold – 2.70-fold; average 1.61-fold). While there remains a few inconsistencies between Fastsimcoal2 and results from our other demographic analysis (e.g. for taxa inhabiting islands south/north of the LGM-sea ice), the amount of increase/decrease was always relatively shallow when occurring in the 'inconsistent' direction (*Supplementary Table 31*). For example, although Fastsimcoal2 inferred a post-LGM population expansion in *E. pachyrhynchus* (where a stable population trajectory was expected), the amount of the increase was only 1.02-fold (compared to *Pygoscelis papua* which experienced a 12.25-fold increase). *Eudyptes filholi* and *E. chrysolophus chrysolophus*/*E. c. schlegeli* were inferred to have experienced a slight N_e decrease following the LGM (1.06-fold and 2.70-fold, respectively), which is similar to the *E. chrysocome* (1.06-fold).

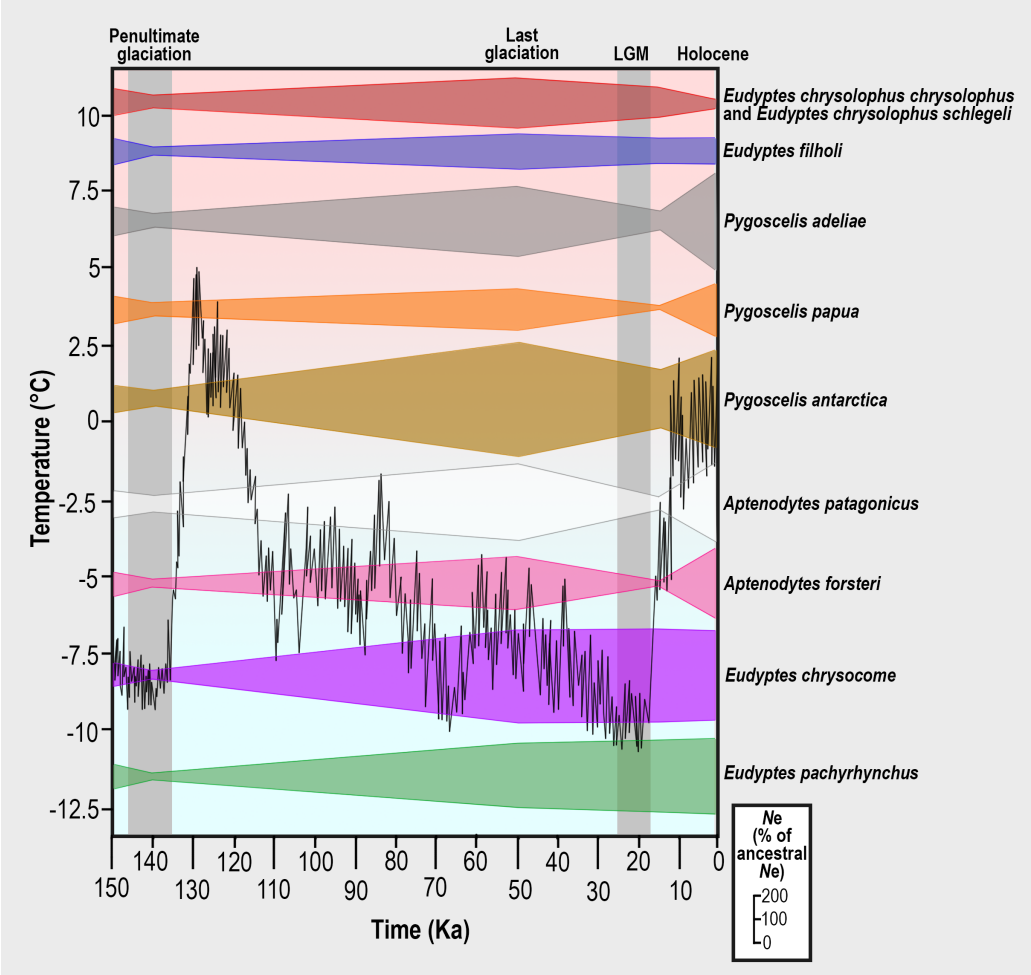


Figure 18. Demographic reconstructions of historical N_e for *Eudyptes*, *Aptenodytes* and *Pygoscelis* penguins inferred by Fastsimcoal2. A 150 Ka record of Antarctic temperature change as estimated from the EPICA Dome Ice Core (Jouzel *et al*., 2007) modified from Cristofari *et al*. (2017). N_e was inferred for five different time periods (mid Pleistocene: 1 mya; Penultimate glaciation: 140 kya; Last glaciation: 50 kya; end of LGM: 15 kya and late Holocene: 1 kya), and is shown as a relative percent of the N_e inferred at 1 Mya, based on the mean values from Fastsimcoal2. Note that N_e at time 1 Mya is shown at 150 Ka. Major climatic periods are indicated. Note also that *E. robustus* and *E. moseleyi* were not included due to small sample size and that *E. chrysolophus chrysolophus* and *E. c. schlegeli* were analysed together.

Tajima's D

While Tajima's D can infer population expansion events, it cannot test for the timing of those expansions. Tajima's D for all penguin taxa were generally consistent with previous demographic results. Past population expansions, inferred by a negative Tajima's D were detected in *Eudyptes chrysolophus chrysolophus*/*E. c. schlegeli*, *E. filholi*, *Pygoscelis adeliae*, *P. antarctica*, *Aptenodytes patagonicus*, *A. forsteri* and possibly *Eudyptes pachyrhynchus*, while stable population sizes were detected in the *Pygoscelis papua*, *Eudyptes moseleyi*, *E. chrysocome* and *E. robustus* (*Supplementary Table 33*). Although Tajima's D was negative for *Eudyptes pachyrhynchus*, which could suggest a population expansion, this result was not statistically significant ($P = 0.135$) (*Supplementary Table 33*).

SNAPP

While the theta value analysis using our SNAPP results did infer some demographic changes (*Supplementary Figure 20*), the internal node ages both within and between penguin taxa vary widely (*Supplementary Figure 21*). For those taxa with older internal node ages (e.g. >600 Ka between some *Eudyptes chrysolophus chrysolophus*/*E. c. schlegeli* populations) the change in theta represents all the fluctuations in N_e that may have occurred over the last 600 Ka. In addition, several species that experienced a post-LGM demographic expansion as inferred by CubSFS, Fastsimcoal2 and Tajima's D; e.g. *Pygoscelis papua*), showed reductions in theta for some sampling localities. However, the internal node ages that at least some of these inferred reductions in theta are based on are far younger than the LGM (see also Clucas *et al*., 2018), increasing the likelihood that contemporary factors impacting N_e are influencing these estimates downward. For these reasons, it is difficult to use these analyses to directly assess the response of N_e to the LGM, and therefore we focus our interpretation on the CubSFS, Fastsimcoal2 and Tajima's D analyses (see above).

Multi-dice

Based on the combined results of CubSFS, Fastsimcoal2, Tajima's D and SNAPP (*Figure 19*), we investigated how synchronous the expansion event was among seven taxa, as well as the timing of the inferred expansion. For our 'expanding' taxa (*Eudyptes chrysolophus chrysolophus*/*E. c. schlegeli*, *E. filholi*, *Pygoscelis adeliae*, *P. papua*, *P. antarctica*, *Aptenodytes patagonicus* and *A. forsteri*), Multi-dice inferred a co-expansion event during the last 10 – 15 Kya while other taxa expanded idiosyncratically (*Figure 19*, *Supplementary Table 34*). Interestingly, our multi-dice results did not reveal a largely synchronous expansion event, as only 28% – 43% of taxa were inferred to co-expand. There are several reasons that may limit

our ability to detect a potentially synchronous expansion event. Mainly, small inaccuracies in the generation times used for each species might lead to large time differences over several thousand of years. Although the 'expanding' taxa might show a general pattern of co-expansion, if the time scale of this co-expansion varied by more than a 1000 years (either due to generation time inaccuracy, or due to real biological signal e.g. *A. forsteri* expanding earlier), then Multi-dice would fail to detect a single co-expansion event.

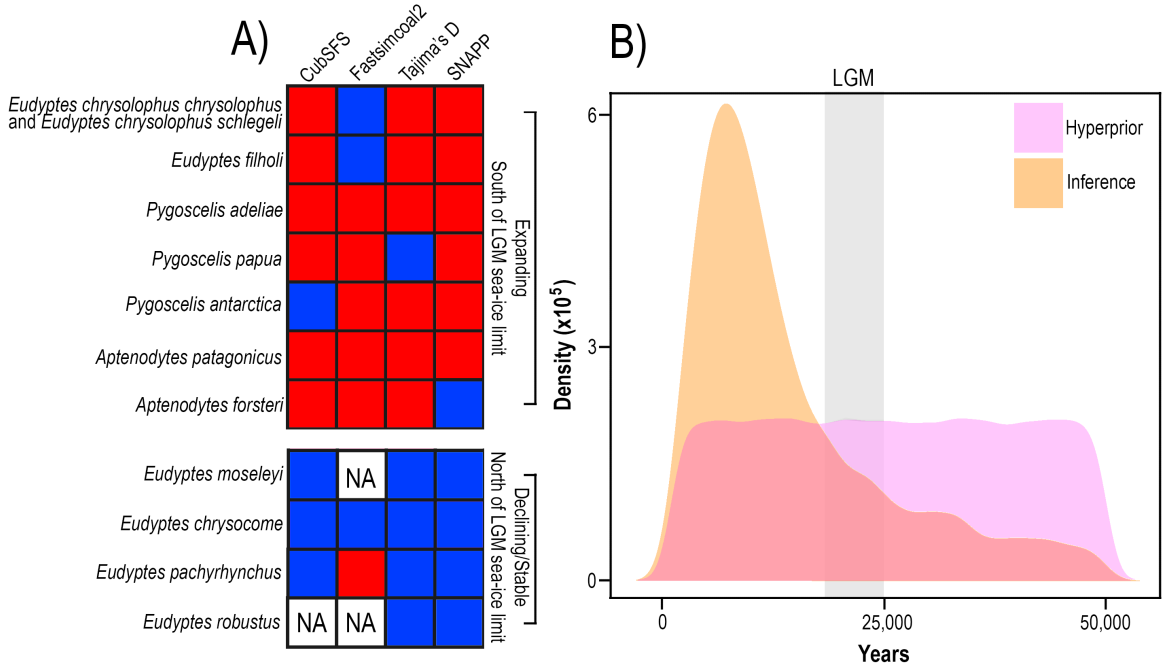


Figure 19. Schematic of demographic results for *Eudyptes*, *Aptenodytes* and *Pygoscelis* penguins. A) shows the combined results of CubSFS, Fastsimcoal2, Tajima's D and SNAPP theta values. Taxa are broadly classified as 'expanding' (red: *Eudyptes chrysolophus chrysolophus*/*E. c. schlegeli*, *E. filholi*, *Pygoscelis adeliae*, *P. papua*, *P. antarctica*, *Aptenodytes patagonicus* and *A. forsteri* [all south of the LGM sea-ice]) or 'declining/stable' (blue: *Eudyptes moseleyi*, *E. chrysocome*, *E. pachyrhynchus* and *E. robustus* [all north of the LGM sea-ice]) on the basis of a majority of these analytical outputs. 'NA' indicates when a species was excluded from a particular analysis due to limited sample size. All analyses specifically address post-LGM demographic change, with the exception of Tajima's D which may also be influenced by earlier demographic events. B) Multi-dice results, suggesting a post-LGM expansion, with the mean of the co-expansion time parameter inferred at 13,342 years (mode: 10,604; median: 10,125). Note that *E. chrysolophus chrysolophus* and *E. c. schlegeli* were analysed together.

Discussion

Parallel post-LGM population expansions detected in eight penguin taxa

Tests for population divergence within the 12 penguin taxa (including previous analyses of *Pygoscelis* and *Aptenodytes* species; see Clucas *et al.*, 2018) revealed little, if any, differentiation (*Figure 14*), with shallow genetic structure consistent with recent demographic and biogeographic expansions following the LGM (*Supplementary Figures 14 – 16*).

Specifically, F_{ST}, PCoA, Structure, DAPC, SNAPP and RAxML revealed shallow genomic structure among isolated sub-Antarctic and Antarctic populations within each taxon (*Figures 14 – 16*, *Supplementary Figures 14 – 16* and *Supplementary Tables 15*, *20 – 29*; see also Clucas *et al*., 2018). In contrast to the recent genetic exchange inferred within most taxa (and between *Eudyptes chrysolophus chrysolophus*/*Eudyptes chrysolophus schlegeli*) these analyses detected little or no introgression among taxa (see *Figures 14 – 16* and *Supplementary Figures 14 – 16*).

We detected strong demographic signals that all penguin taxa breeding south of the LGM sea-ice zone (*E. chrysolophus chrysolophus*/*E. c. schlegeli*, *E. filholi*, *Aptenodytes patagonicus*, *A. forsteri*, *Pygoscelis adeliae*, *P. papua*, *P. antarctica*) experienced post-LGM demographic expansions, while four taxa breeding north of the LGM sea-ice zone (*Eudyptes moseleyi*, *E. chrysocome*, *E. pachyrhynchus* and possibly *E. robustus*) have experienced relatively stable demographic histories (*Figures 17 – 19, Supplementary Figures 17 – 21* and *Supplementary Tables 30 – 34*). Crucially, our demographic analyses of CubSFS, Fastsimcoal2, Tajima's D and SNAPP theta values produced consistent signals for a post-LGM multi-taxon expansion event that followed the reduction of LGM winter sea-ice (*Figure 19*). We detected some variation in the outcomes of individual demographic analyses for particular species (*Figure 19*), perhaps a reflection of varying sensitivity of different model-based approaches. However, a combination of results across multiple analyses revealed consistently 'stable/declining' demographic trajectories for taxa inhabiting LGM ice-free regions, versus consistently 'expanding' trajectories for LGM ice-affected taxa (*Figure 19*). Reduction of sea-ice therefore likely provided new opportunities for taxa to recolonise the southern regions.

LGM sea-ice restricted breeding success

With the exception of the ice-adapted *Aptenodytes forsteri* (which breeds on the sea-ice around coastal Antarctica), all penguin taxa are reliant on ice-free terrain to build nests during the breeding season, while all require access to the open ocean to forage. Extensive glaciation across the high-latitudes during the LGM would thus have become a major barrier to successful breeding of penguins, while extensive sea-ice would have restricted foraging access for all southern populations. Past studies have found some evidence to suggest there were persistent polynyas during the LGM in regions south of the LGM sea-ice zone (Emslie *et al*., 2018). Localised refugial populations of some of these southern taxa close to such polynyas may have been able to persist in the high-latitudes during the LGM (Brambati *et al*., 2002; Thatje *et al*., 2008; Younger *et al*., 2015c; Clucas *et al*., 2016). However, the majority of taxa south of the LGM sea-ice would likely have become locally extinct.

Dispersal facilitates biogeographic shifts

Although LGM breeding ranges for Southern Ocean penguins remain unclear, our results may support previous hypotheses of Fraser *et al.* (2011), that during the LGM, sub-Antarctic and Antarctic taxa retreated to ice-free refugia such as Gough, Amsterdam, and the Falkland Islands, southern South America, and NZ's southern islands (see *Figure 20*; and Fraser *et al.*, 2009; Fraser *et al.*, 2011; Rains *et al.*, 2011; Fraser *et al.*, 2018). Post-LGM reductions in sea-ice (*Figure 20*) subsequently transformed Southern Ocean coastal ecosystems (Fraser *et al.*, 2009; Trucchi *et al.*, 2014), facilitating rapid re-colonisation of high-latitude sub-Antarctic and Antarctic shores, perhaps following the eastward path of the ACC (Fraser *et al.*, 2009; Fraser *et al.*, 2018). While LGM coastal regions, and subfossil remains are now largely underwater (see *Figure 20*), potential refugial regions are suggested by current distributions (e.g. *E. filholi* likely expanded south from glacial refugia in NZ's ice-free Auckland, Campbell, and Antipodes islands; *Figures 14, 20* and *Supplementary Figure 13*). While previous studies have suggested that the Southern Ocean's circumpolar fronts (STF and APF) represent barriers to dispersal for many marine species (Frugone *et al.*, 2018), including seabirds (Friesen, 2005; Friesen *et al.*, 2007; Munro & Burg, 2017) penguins commonly traverse these major oceanographic boundaries (Thiébot *et al.*, 2011; Thiébot *et al.*, 2012). This exceptional dispersal ability thus provides a key mechanism facilitating rapid biogeographic and demographic shifts (e.g. retractions and expansions) in response to changing climate (see also de Bruyn *et al.*, 2009).

Stable/declining populations in four penguin species

Our results revealed consistently 'stable/declining' demographic trajectories for all studied penguin taxa inhabiting regions north of the LGM sea-ice limit (*Eudyptes pachyrhynchus*, *E. robustus*, *E. chrysocome* and *E. moseleyi*). While there is some terrestrial evidence for LGM ice expansion on the NZ Auckland and Campbell Islands (Hodgson *et al.*, 2014), this ice extent is unlikely to have restricted breeding and foraging of those penguin populations. The clear pattern of demographic expansions in taxa south of the LGM sea-ice limit compared to stable/declining demographic histories in taxa north of the LGM sea-ice limit is concordant with the biological requirements of penguins. As such, taxa breeding north of the LGM sea-ice limit were not forced into refugia. Moreover, during the LGM, sea-levels were 120 – 135 m lower than present day (Clark & Mix, 2002), which may have created additional opportunities for the establishment of large breeding colonies of penguins in these northern regions. As a consequence, rising sea-levels following the LGM would likely have led to the local extinction of many of these colonies, which is potentially corroborated by the post-LGM population declines inferred for many of these northern penguin taxa.

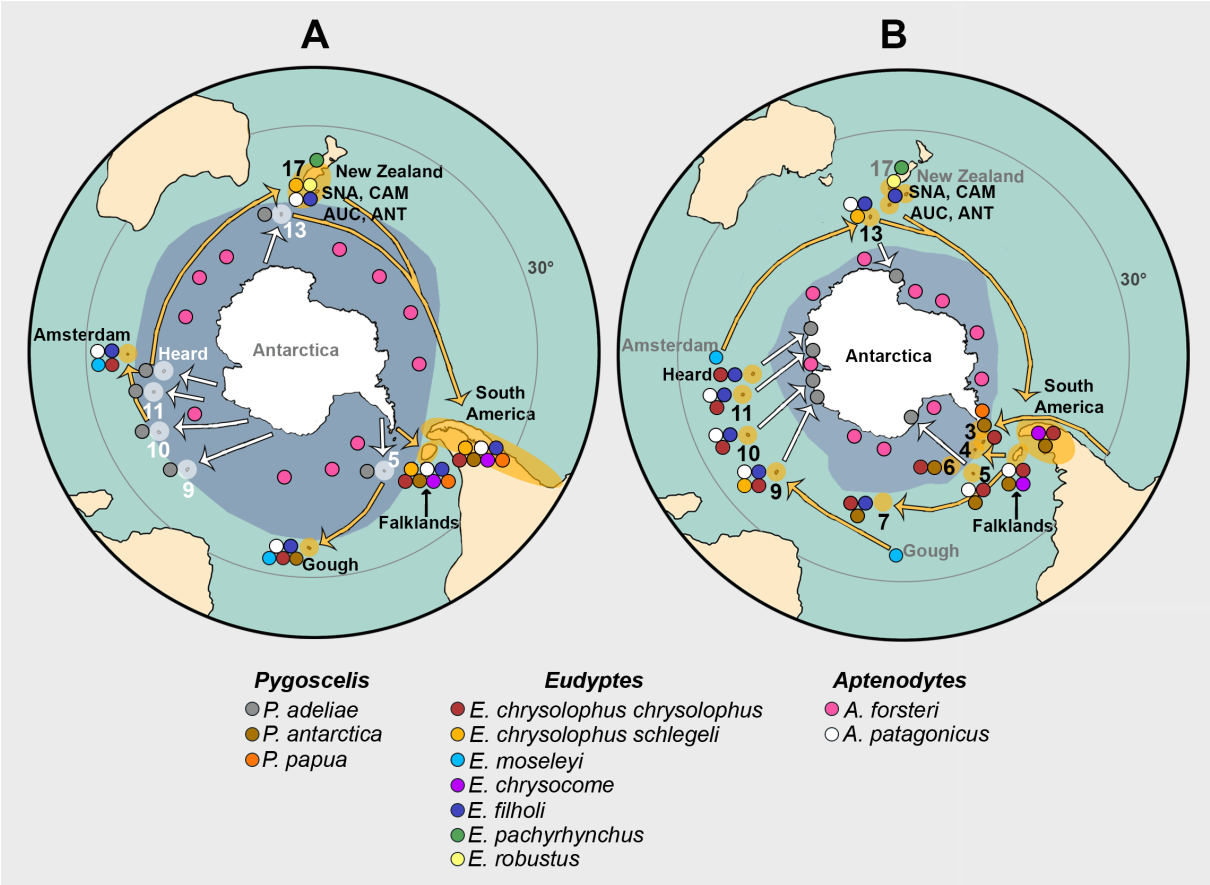


Figure 20. Population expansion and contractions for *Eudyptes*, *Aptenodytes* and *Pygoscelis* penguins in relation to post-LGM climate change. **A)** shows the winter sea-ice and sea-level during the LGM, with putative the refugia shown (orange ellipses for sub-Antarctic penguins; grey ellipses for all Antarctic penguins except *Aptenodytes forsteri*), the routes of refugia is shown: yellow arrows for sub-Antarctic species; white arrows for Antarctic species. *A. forsteri* presumably bred on the fringes of the summer sea-ice during the LGM (indicated by the pink dots). Site names in black indicate possible refugial regions for sub-Antarctic penguins, while sites encircled in opaque white indicate possible refugial regions for Antarctic penguins. **B)** shows the present sea-level and winter sea-ice extent, with possible post-LGM routes of recolonisation into sub-Antarctic and Antarctic habitats (white arrows for Antarctic penguins, except *A. forsteri*), yellow arrows for sub-Antarctic species. *A. forsteri* is again marked with pink dots. Site names in black indicate where each penguin currently breeds, while site names in grey indicate locations where those taxa may have bred during the LGM (as shown in A]). Note, these LGM breeding ranges are uncertain. The maps have been adapted from Fraser *et al*. (2011).

Contrasting patterns among Southern Ocean taxa

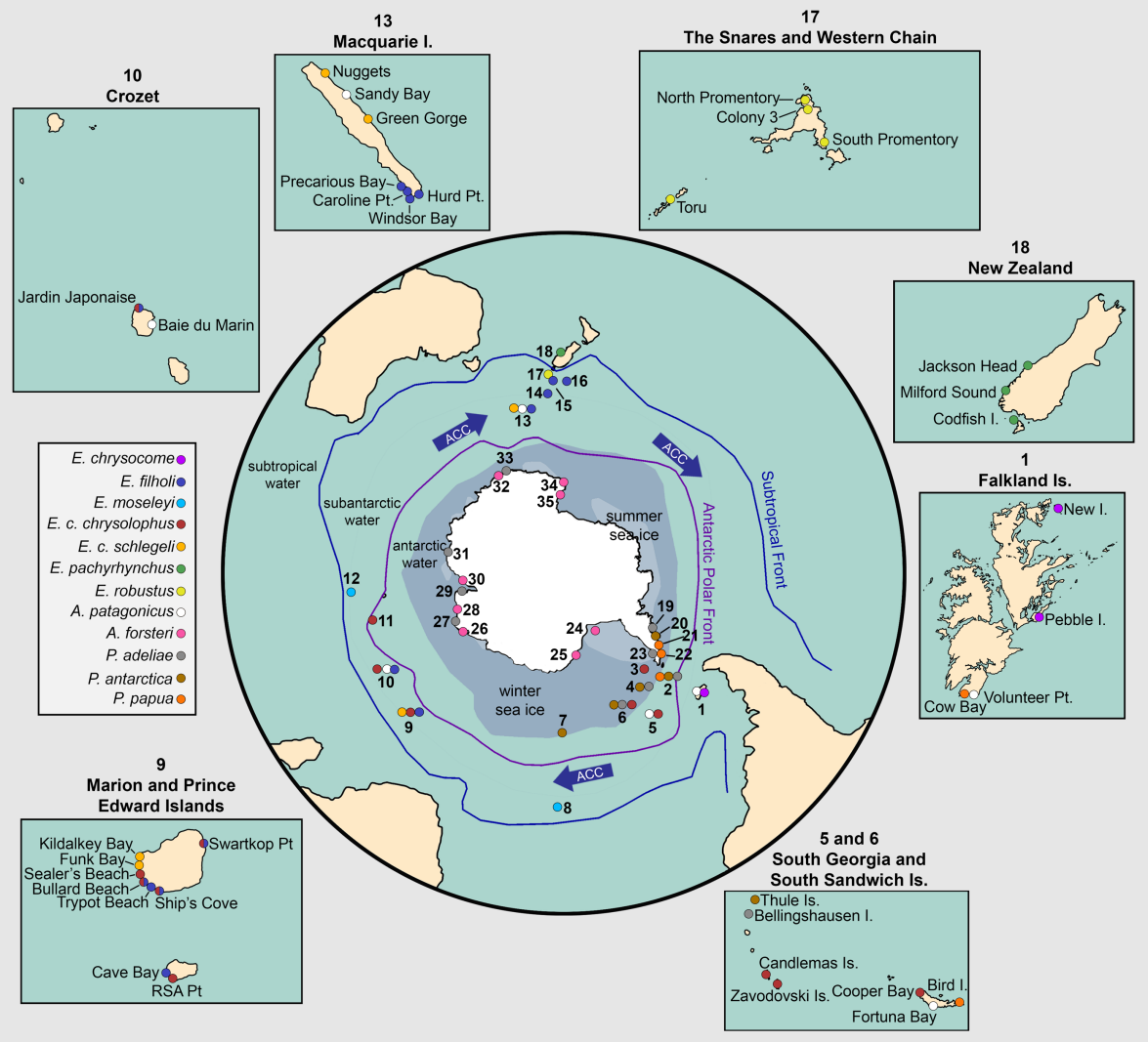
Glacial cycles can have profound effects on the biogeographic and demographic structure of high-latitude taxa. While our demographic analyses demonstrate rapid post-LGM expansions in eight penguin taxa, other studies have found a lack of population expansions following Southern Ocean Holocene climate warming. For example, demographic reconstructions for the snow petrel (*Pagodroma nivea*) suggests that at least two populations may have persisted in ice-free regions during the LGM (Carrea *et al*., 2019), with only one population experiencing a post-LGM expansion. The authors suggest that the ability to fly may represent an advantage for *Pagodroma nivea* over flightless taxa in persisting during the LGM. Demographic reconstructions for the weddell seal (*Leptonychotes weddellii*) suggest that their populations

have remained stable for at least 75 thousand years (Younger *et al*., 2016). In contrast, demographic reconstructions and population structure analyses of the southern elephant seal (*Mirounga leonina*), *Azorella* plants (*A. selago* and *A. macquariensis*) and *Durvillaea antarctica* have revealed rapid population expansions following post-LGM recolonization of founder individuals (de Bruyn *et al*., 2009; Fraser *et al*., 2009; Corrigan *et al*., 2016; Chau *et al*., 2019; see also *Figure 3*). The choice of mutation rate, and possibly time-dependency issues in the modelling approaches may play some part in the conflicting patterns inferred for Southern Ocean taxa. The contrasting responses to post-LGM climate change across high-latitude taxa may also suggest that differing behaviour (e.g. philopatry) adaptive capacities (e.g. dispersal ability) and/or fine-scale niche preferences (e.g. foraging locations) may play an important role in a species response to climate change.

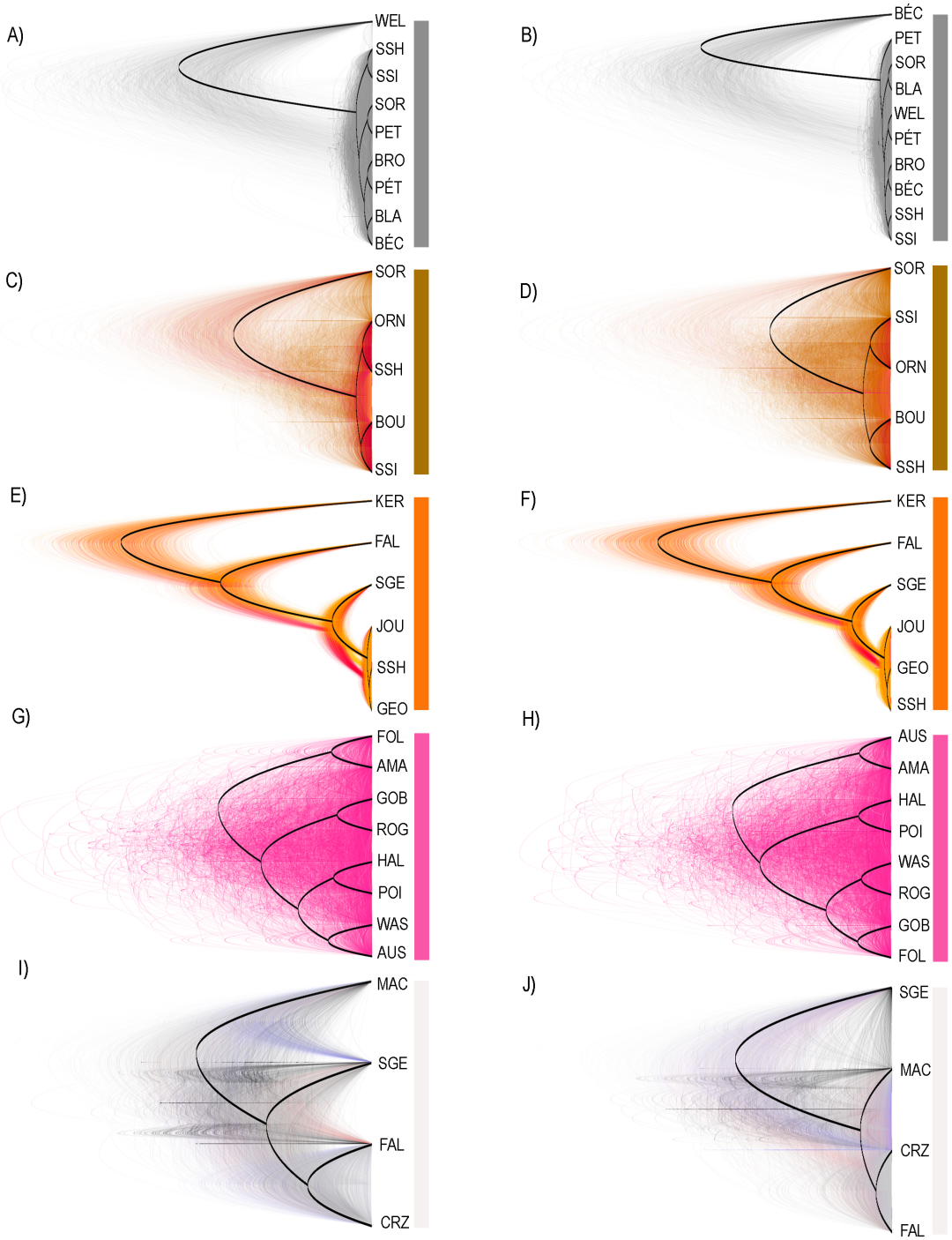
Conclusions

Understanding how biota responded to past global climate change is essential for predicting species distributions and population sizes under future climate projections and for developing appropriate conservation management strategies (Forcada & Trathan, 2009; Dussex *et al*., 2019). While global temperatures continue to increase, terrestrial and coastal biota from mid-latitude habitats will continue to shift towards the poles (Fraser *et al*., 2018; see also Jagatap *et al*., 2019) or alternatively face extinction, as they are pushed to their southernmost limits (Fraser *et al*., 2011; Fraser *et al*., 2018). Many penguin populations are declining, or are predicted to decline as warming continues (Barbraud & Weimerskirch, 2001; Le Bohec *et al*., 2008). *Pygoscelis adeliae* and *Aptenodytes forsteri* populations, for example, have declined or even disappeared at some of their northernmost localities around Antarctica (Barbraud & Weimerskirch, 2001; Forcada *et al*., 2006). In the case of *A. forsteri* these declines have been linked directly to reductions in sea-ice (Fretwell & Trathan, 2019). By contrast, populations of *Pygoscelis papua* are apparently expanding their ranges southward as the climate warms (Lynch *et al*., 2012). Our study demonstrates the sensitivity of demography and distribution of keystone species within Southern Ocean ecosystems to the effects of climate change, and reveals the potential for major future biological shifts in this rapidly warming part of the planet.

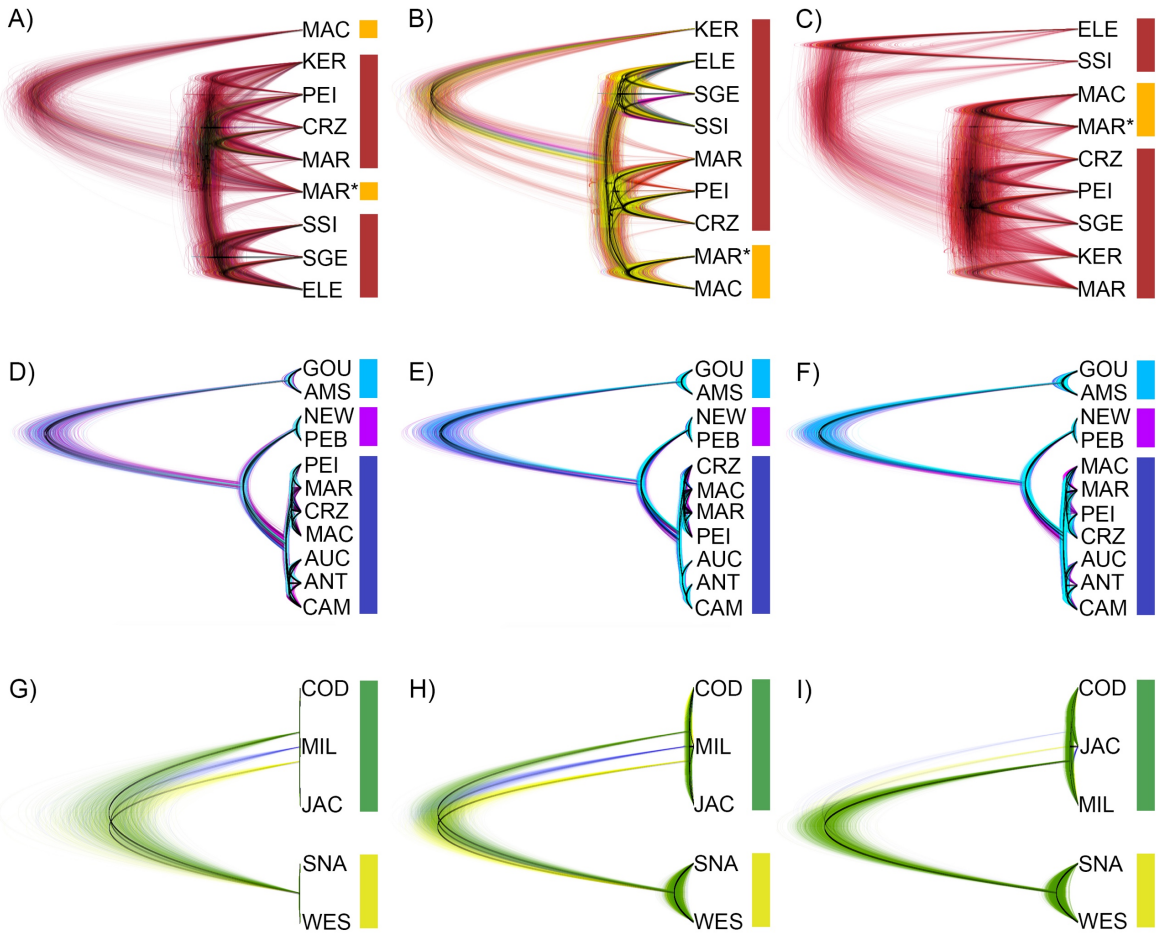
Supplementary Figures for *Chapter 3*



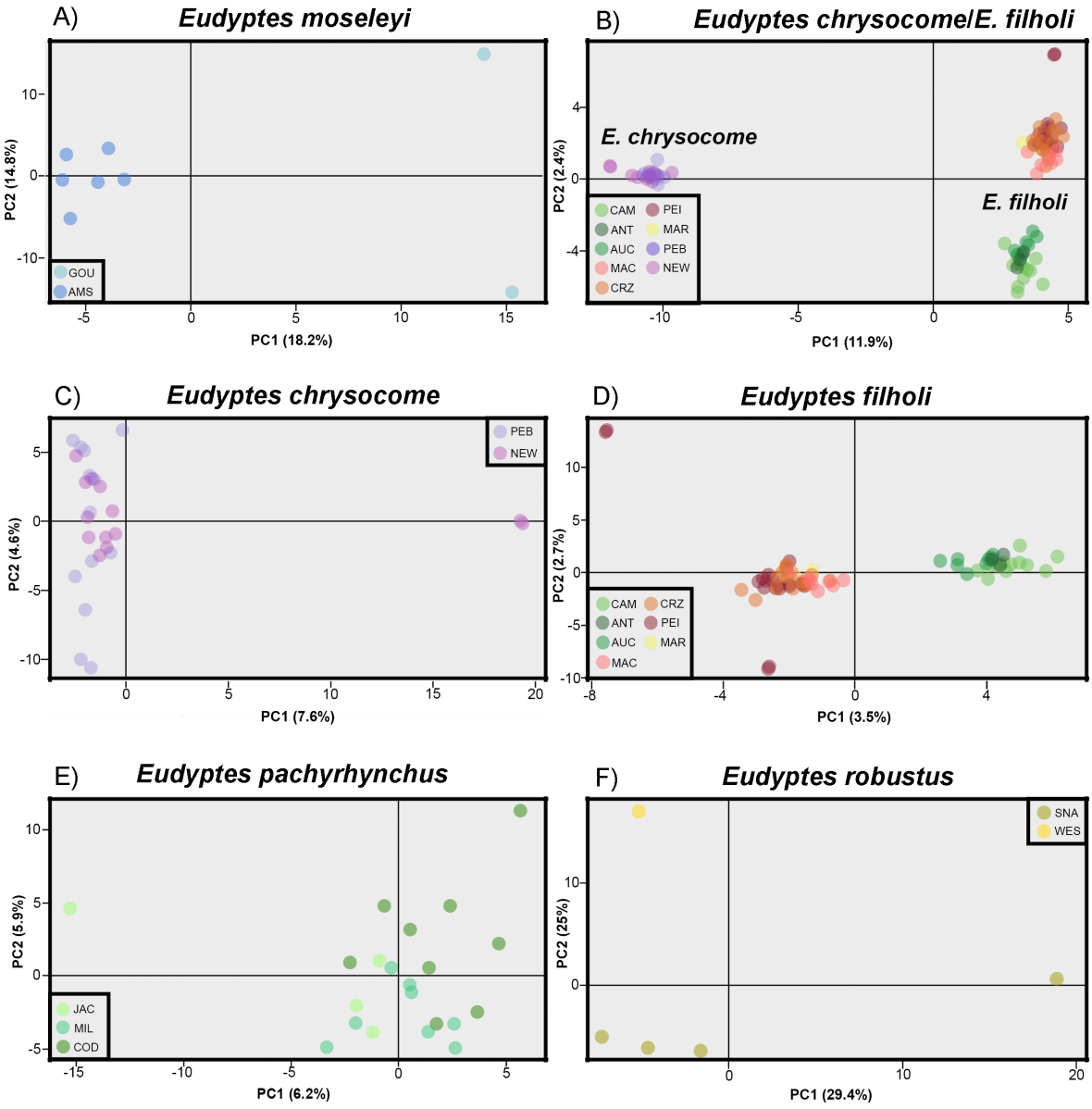
Supplementary Figure 13. Detailed map of sampling locations for *Eudyptes*, *Aptenodytes* and *Pygoscelis* penguins. Island groups that contained multiple sample collections from different sites are shown. Note *P. adelie* and *P. antarctica* South Orkney Islands samples all came from Signy Island, and *P. adelie*, *P. antarctica* and *P. papua* South Shetland Islands samples all came from Admiralty Bay and *P. papua* Kerguelen samples came from Pointe du Marne (see *Supplementary Table 16*). The main map has been modified from Fraser *et al*. (2011) and shows current winter and summer sea-ice, the STF, the APF and the ACC. Site details and codes used in other figures and tables are follows: 1 is Falkland Islands (PEB, NEW and FAL); 2 is South Sandwich Islands (SSH); 3 is Elephant Island (ELE); 4 is South Orkney Islands (SOR); 5 is South Georgia (SGE); 6 is South Sandwich Islands (SSI); 7 is Bouvet (BOU); 8 is Gough Island and Tristan da Cunha (GOU); 9 is Marion Island and Prince Edward Islands (MAR and PEI); 10 is Crozet (CRZ); 11 is Kerguelen (KER); 12 is Amsterdam Island (AMS); 13 is Macquarie Island (MAC); 14 is Campbell Island (CAM); 15 is Auckland Islands (AUC); 16 is Antipodes Islands (ANT); 17 is The Snares and Western Chain (SNA and WES); 18 is Codfish Island, Milford Sound and Jackson Head (COD, MIL and JAC); 19 is Peterman Island (PET); 20 is Orne Harbour (ORN); 21 is Jougla Point (JOU); 22 is George's Point (GEO); 23 is Brown Bluff (BRO); 24 is Gould Bay (GOB); 25 is Halley Bay (HAL); 26 is Fold Island (FOL); 27 is Béchervaise Island (BÉC); 28 is Auster (AUS); 29 is Welch Island (WEL); 30 is Amanda Bay (AMA); 31 is Blakeney Point (BLA); 32 is Point Géologie (POI); 33 is Pétrels Island (PÉT); 34 is Cape Roget (ROG) and 35 is Cape Washington (WAS).



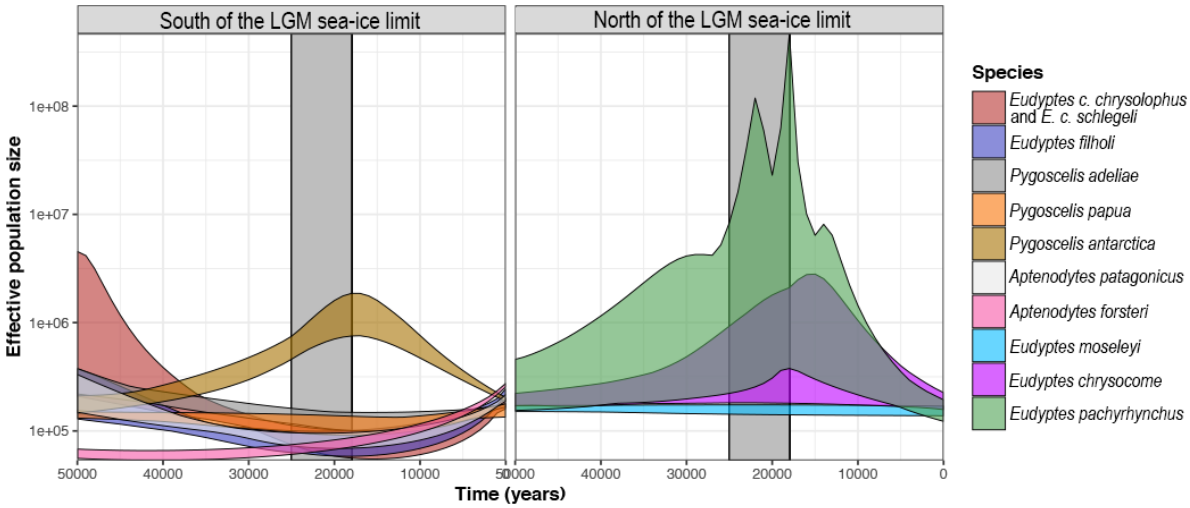
Supplementary Figure 14. Alternative SNAPP analyses for *Pygoscelis* and *Aptenodytes* penguins. **(A) – (B)** is *Pygoscelis adeliae* (grey); **(C) – (D)** is *P. antarctica* (brown); **(E) – (F)** is *P. papua* (orange); **(G) – (H)** is *Aptenodytes forsteri* (pink); and **(I) – (J)** is *A. patagonicus* (white). Note the SNAPP analyses for *Pygoscelis papua* is taken from Clucas *et al*. (2018), the SNAPP analysis for *Aptenodytes forsteri* is taken from Younger *et al*. (2017) and the SNAPP analysis for *A. patagonicus* is taken from Clucas *et al*. (2016). Colours have been modified for this study. The SNAPP analyses for *Pygoscelis adeliae* and *P. antarctica* are new. Site details are follows: WEL is Welch Island; SSH is South Shetland Islands; SSI is South Sandwich Islands; SOR is South Orkney Islands; PET is Peterman Island; BRO is Brown Bluff; PÉT is Pétrels Island; BLA is Blakeney Point; BÉC is Béchervaise Island; ORN is Orne Harbour; BOU is Bouvet; KER is Kerguelen; FAL is Falkland Islands; SGE is South Georgia; JOU is Jougla Point; GEO is George's Point; FOL is Fold Island; AMA is Amanda Bay; GOB is Gould Bay; ROG is Cape Roget; HAL is Halley Bay; POI is Point Géologie; WAS is Cape Washington; AUS is Auster; MAC is Macquarie Island; and CRZ is Crozet.



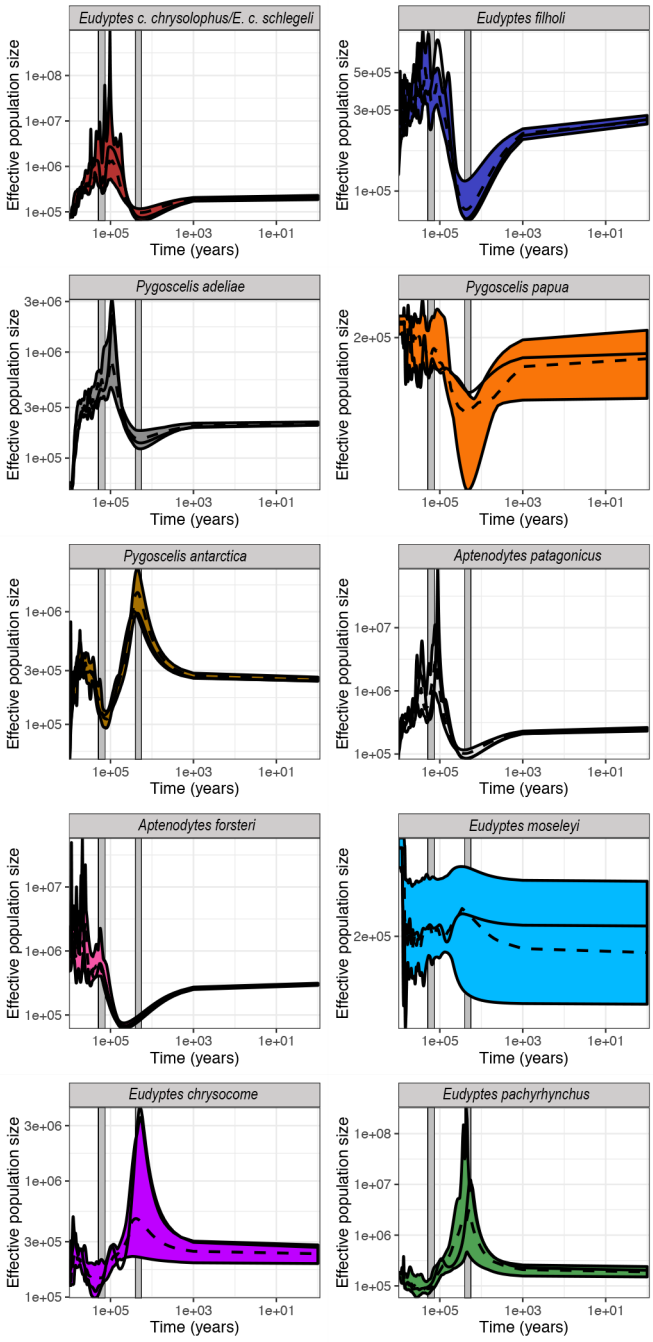
Supplementary Figure 15. Alternative SNAPP analyses for closely related *Eudyptes* penguins. **(A) – (C)** is *Eudyptes chrysolophus chrysolophus* (red) and *E. c. schlegeli* (orange); **(D) – (F)** is *E. moseleyi* (light blue), *E. chrysocome* (purple) and *E. filholi* (dark blue); **(G) – (I)** is *E. pachyrhynchus* (green) and *E. robustus* (yellow). MAR* indicates Marion Island 'white faced' penguins of *E. c. chrysolophus*/*E. c. schlegeli*. ELE is Elephant Island; SGE is South Georgia; SSI is South Sandwich Islands; MAR is Marion Island; PEI is Prince Edward Islands; CRZ is Crozet; KER is Kerguelen; MAR* is Marion Island 'white faced' penguins; MAC is Macquarie Island; AMS is Amsterdam Island; GOU is Gough Island; NEW is New Island; PEB is Pebble Beach; CAM is Campbell Island; AUC is Auckland Islands; ANT is Antipodes Islands; COD is Codfish Island; MIL is Milford Sound; JAC is Jackson Head; SNA is the Snares; and WES is the Western Chain.



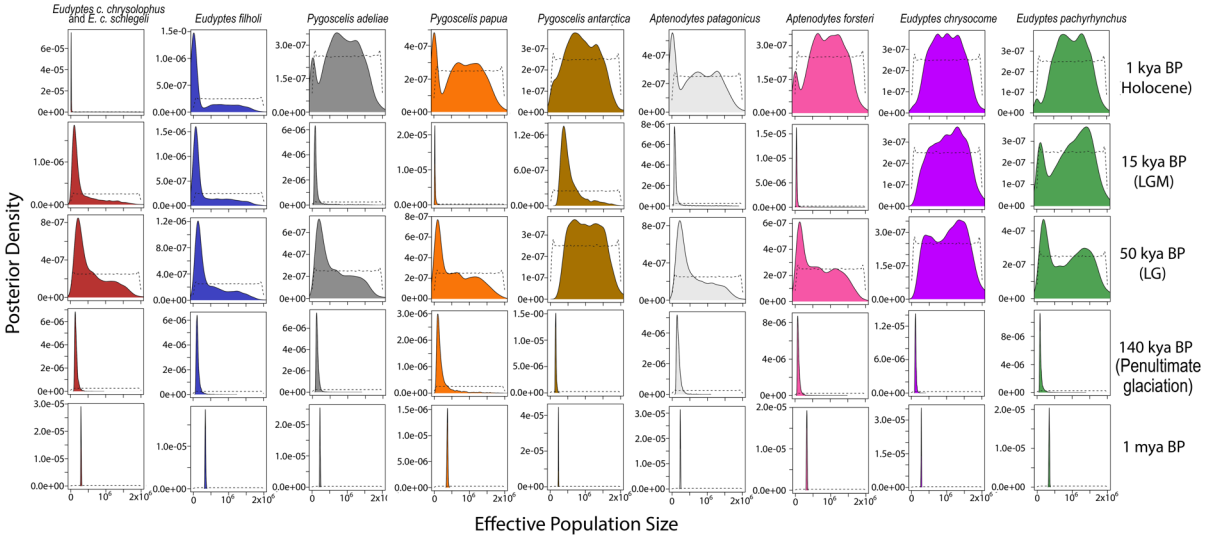
Supplementary Figure 16. PCoA of within and between species genetic variation for closely-related *Eudyptes* penguins. **(A)** is *E. moseleyi*; **(B)** is *Eudyptes chrysocome* and *E. filholi*; **(C)** is *E. chrysocome*; **(D)** is *E. filholi*; **(E)** is *E. pachyrhynchus*; and **(F)** is *E. robustus*. Additional PCoAs of between species genetic variation (except for *Eudyptes chrysocome* and *E. filholi*, which is shown here) are shown in *Figure 16*. GOU is Gough Island; AMS is Amsterdam Island; CAM is Campbell Island; ANT is Antipodes Islands; AUC is Auckland Islands; MAC is Macquarie Island; CRZ is Crozet; PEI is Prince Edward Island; MAR is Marion Island; PEB is Pebble Beach; NEW is New Island; JAC is Jackson Head; MIL is Milford Sound; COD is Codfish Island; SNA is The Snares; and WES is Western Chain.



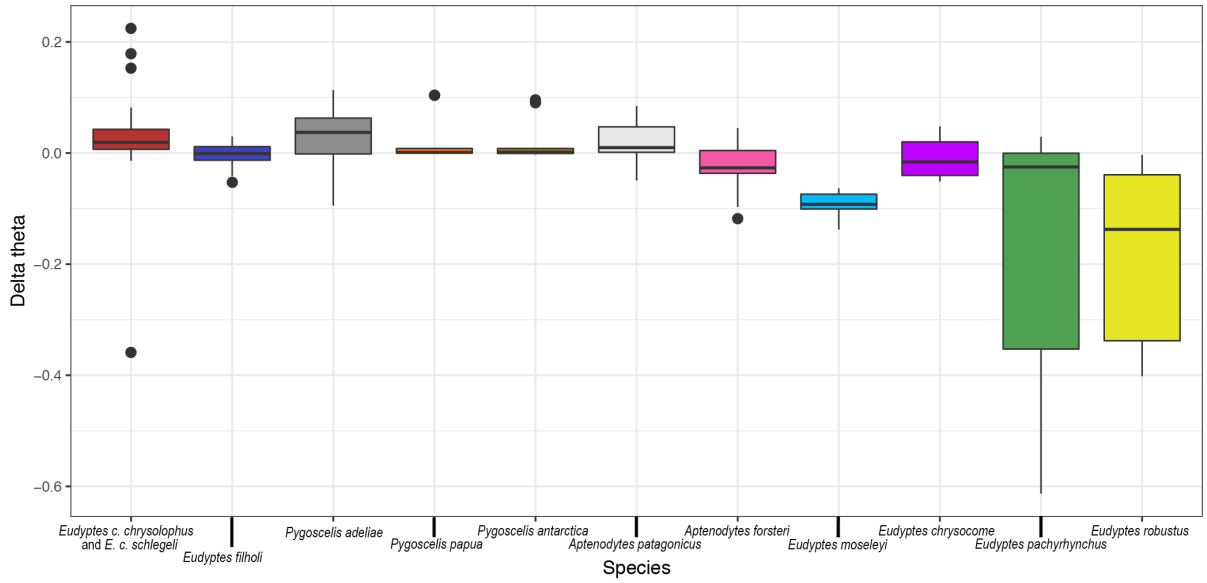
Supplementary Figure 17. Overlaid 95% CIs for *Eudyptes*, *Pygoscelis* and *Aptenodytes* penguins north and south of the LGM sea-ice limit based on CubSFS demographic reconstructions. Note the extremely wide confidence interval for *Eudyptes pachyrhynchus*, and two broad categories of demographic reconstruction; (1) recently expanding populations since the LGM (*E. chrysolophus chrysolophus*/*E. c. schlegeli*, *E. filholi*, *Pygoscelis papua*, *P. adeliae*, *Aptenodytes patagonicus* and *A. forsteri*); or (2) constant/declining populations since the LGM (*Eudyptes pachyrhynchus*, *E. chrysocome*, *E. moseleyi* and *Pygoscelis antarctica*). The solid grey bar represents the LGM. Note that *Eudyptes chrysolophus chrysolophus* and *E. c. schlegeli* are combined in the same analysis. *E. robustus* was not analysed due to low sample size. Refer to *Figure 17* and *Supplementary Figure 18* for each species plotted separately.



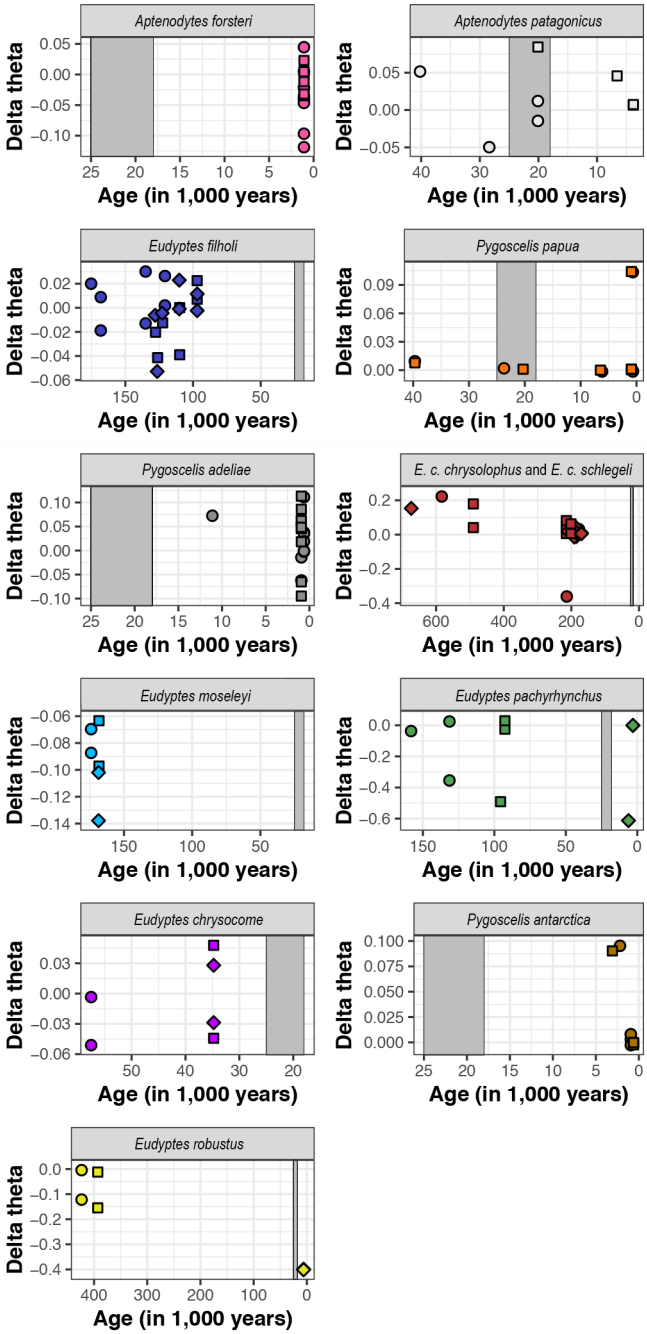
Supplementary Figure 18. CubSFS demographic reconstructions for *Eudyptes*, *Pygoscelis* and *Aptenodytes* penguins spanning 1 Ma. 95% confidence intervals are given by solid colour intervals. Median for bootstrap replicates is given by the dotted line, and the solid line gives the demographic reconstruction for the amended SFS. Note that *Eudyptes chrysolophus chrysolophus* and *E. c. schlegeli* are combined in the same analysis. *E. robustus* was not analysed due to low sample size. The left grey bar indicates the Penultimate Glaciation (135 kya) and the right grey bar indicates the LGM (18 – 25 kya).



Supplementary Figure 19. Fastsimcoal2 posterior distributions of N_e for *Eudyptes*, *Pygoscelis* and *Aptenodytes* at five different time points spanning 1 million years. 1 kya BP (Holocene), 15 kya BP (LGM), 50 kya BP (Last Glacial), 140 kya BP (Penultimate Glaciation) and 1 mya BP. Dashed line represents prior distributions. Note that *Eudyptes chrysolophus chrysolophus*/*E. c. schlegeli* are considered a single taxon. *E. robustus* and *E. moseleyi* were not analysed due to low sample size



Supplementary Figure 20. SNAPP delta theta values for *Eudyptes*, *Pygoscelis* and *Aptenodytes* penguins, based on delta theta of individual sampled locations. Positive theta values indicate expansions, negative, population contractions. Taxa that are inferred to have undergone a population expansion include *Eudyptes chrysolophus chrysolophus*/*E. c. schlegeli*, *E. filholi*, *Pygoscelis adeliae*, *P. papua*, *P. antarctica* and *Aptenodytes patagonicus*. Taxa that are not inferred to have undergone a population expansion following the LGM include *Eudyptes moseleyi*, *E. chrysocome*, *E. pachyrhynchus*, *E. robustus* and *Aptenodytes forsteri*). Black dots represent theta inferences of locations that are more than 1.5 X the size of the inter-quartile range above or below the upper or lower quartile, respectively.



Supplementary Figure 21. SNAPP delta theta values for *Eudyptes*, *Pygoscelis* and *Aptenodytes* penguins in relation to node age. Age is taken from the node ages in a number of substitutions from the annotated tree files, divided by the mutation rate and multiplied by the taxon-specific generation time. Circles are replicate one (*Eudyptes*, *Pygoscelis* and *Aptenodytes*), squares are replicate two (*Eudyptes*, *Pygoscelis* and *Aptenodytes*) and diamonds are replicate three (*Eudyptes* only). The solid grey bar indicates the LGM. Note all *Aptenodytes forsteri*, *Pygoscelis adeliae* and *P. antarctica* penguins have post-LGM node ages.

Supplementary Table 14. *Eudyptes* samples used for obtaining genome-wide SNPs. DNA samples in category 1 were too low quality for DArT-seq™ library preparation, DNA samples in category 2 were sent for DArT-seq™ (but failed quality control), and DNA samples in category 3 were sequenced and genotyped by DArT-seq™. DNA samples that were extracted several times are indicated with multiple sample codes. Island code can be directly compared to other tables and figures in this study.

Sample Code	Taxon	Location	Island Group	Island Code	Sample Type	Date Collected	Sample Identifier	Category
JW718	*Eudyptes chrysocome*	New Island	Falkland Islands	1, NEW	Blood	01/2010	370 M	1
JW719	*Eudyptes chrysocome*	New Island	Falkland Islands	1, NEW	Blood	01/2010	382 F	1
JW720	*Eudyptes chrysocome*	New Island	Falkland Islands	1, NEW	Blood	01/2010	374 M	3
JW721	*Eudyptes chrysocome*	New Island	Falkland Islands	1, NEW	Blood	01/2010	391 F	3
JW722	*Eudyptes chrysocome*	New Island	Falkland Islands	1, NEW	Blood	01/2010	392 M	3
JW723	*Eudyptes chrysocome*	New Island	Falkland Islands	1, NEW	Blood	01/2010	384 F	1
JW724	*Eudyptes chrysocome*	New Island	Falkland Islands	1, NEW	Blood	01/2010	365 M	3
JW725	*Eudyptes chrysocome*	New Island	Falkland Islands	1, NEW	Blood	01/2010	394 F	1
JW726	*Eudyptes chrysocome*	New Island	Falkland Islands	1, NEW	Blood	01/2010	396 F	3
JW727	*Eudyptes chrysocome*	New Island	Falkland Islands	1, NEW	Blood	01/2010	367 F	1
JW728	*Eudyptes chrysocome*	New Island	Falkland Islands	1, NEW	Blood	01/2010	373 F	3
JW729	*Eudyptes chrysocome*	New Island	Falkland Islands	1, NEW	Blood	01/2010	368 F	1
JW730	*Eudyptes chrysocome*	New Island	Falkland Islands	1, NEW	Blood	01/2010	366 M	3
JW731	*Eudyptes chrysocome*	New Island	Falkland Islands	1, NEW	Blood	01/2010	378 M	3
JW732	*Eudyptes chrysocome*	New Island	Falkland Islands	1, NEW	Blood	01/2010	377 F	3
JW733	*Eudyptes chrysocome*	New Island	Falkland Islands	1, NEW	Blood	01/2010	364 M	3
JW734	*Eudyptes chrysocome*	New Island	Falkland Islands	1, NEW	Blood	01/2010	369 F	3
JW735	*Eudyptes chrysocome*	New Island	Falkland Islands	1, NEW	Blood	01/2010	376 F	3
JW736	*Eudyptes chrysocome*	New Island	Falkland Islands	1, NEW	Blood	01/2010	379 M	1
JW737	*Eudyptes chrysocome*	New Island	Falkland Islands	1, NEW	Blood	01/2010	D036	3
JW738	*Eudyptes chrysocome*	New Island	Falkland Islands	1, NEW	Blood	01/2010	375 F	1
JW739	*Eudyptes chrysocome*	New Island	Falkland Islands	1, NEW	Blood	01/2010	371 F	1
JW740	*Eudyptes chrysocome*	New Island	Falkland Islands	1, NEW	Blood	01/2010	385 M	1
JW1080	*Eudyptes chrysocome*	Pebble Beach	Falkland Islands	1, PEB	Blood	12/2006	R9	3
JW1081	*Eudyptes chrysocome*	Pebble Beach	Falkland Islands	1, PEB	Blood	12/2006	R10	3
JW1082	*Eudyptes chrysocome*	Pebble Beach	Falkland Islands	1, PEB	Blood	12/2006	R11	3
JW1083	*Eudyptes chrysocome*	Pebble Beach	Falkland Islands	1, PEB	Blood	12/2006	R12	3
JW1084	*Eudyptes chrysocome*	Pebble Beach	Falkland Islands	1, PEB	Blood	12/2006	R13	3
JW1085	*Eudyptes chrysocome*	Pebble Beach	Falkland Islands	1, PEB	Blood	12/2006	R14	3
JW1086	*Eudyptes chrysocome*	Pebble Beach	Falkland Islands	1, PEB	Blood	12/2006	R15	3
JW1087	*Eudyptes chrysocome*	Pebble Beach	Falkland Islands	1, PEB	Blood	12/2006	R16	3

Sample Code	Taxon	Location	Island Group	Island Code	Sample Type	Date Collected	Sample Identifier	Category
JW1088	*Eudyptes chrysocome*	Pebble Beach	Falkland Islands	1, PEB	Blood	12/2006	R17	1
JW1089	*Eudyptes chrysocome*	Pebble Beach	Falkland Islands	1, PEB	Blood	12/2006	R18	3
JW1090	*Eudyptes chrysocome*	Pebble Beach	Falkland Islands	1, PEB	Blood	12/2006	R19	3
JW1091	*Eudyptes chrysocome*	Pebble Beach	Falkland Islands	1, PEB	Blood	12/2006	R20	1
JW1092	*Eudyptes chrysocome*	Pebble Beach	Falkland Islands	1, PEB	Blood	12/2006	R21	3
JW1093	*Eudyptes chrysocome*	Pebble Beach	Falkland Islands	1, PEB	Blood	12/2006	R22	1
JW1094	*Eudyptes chrysocome*	Pebble Beach	Falkland Islands	1, PEB	Blood	12/2006	R23	3
JW1095	*Eudyptes chrysocome*	Pebble Beach	Falkland Islands	1, PEB	Blood	12/2006	R24	1
JW1096	*Eudyptes chrysocome*	Pebble Beach	Falkland Islands	1, PEB	Blood	12/2006	R25	3
JW1097	*Eudyptes chrysocome*	Pebble Beach	Falkland Islands	1, PEB	Blood	12/2006	R26	1
JW1098	*Eudyptes chrysocome*	Pebble Beach	Falkland Islands	1, PEB	Blood	12/2006	R27	1
JW1099	*Eudyptes chrysocome*	Pebble Beach	Falkland Islands	1, PEB	Blood	12/2006	R28	1
JW526	*Eudyptes filholi*	Campbell Island	Campbell Island	14, CAM	Blood	23/10/2013	PIT 153746471 GLS #018 Male	3
JW527	*Eudyptes filholi*	Campbell Island	Campbell Island	14, CAM	Blood	23/10/2013	0153723532 GLS #010 Male	1
JW528	*Eudyptes filholi*	Campbell Island	Campbell Island	14, CAM	Blood	10/11/2013	167856729 GLS #8 Female	3
JW529	*Eudyptes filholi*	Campbell Island	Campbell Island	14, CAM	Blood	23/10/2013	Pit Tag 167887471 GLS #045 Female	2
JW530, JW568	*Eudyptes filholi*	Campbell Island	Campbell Island	14, CAM	Blood	23/10/2013	167887648 GLS #6 Female	1
JW531	*Eudyptes filholi*	Campbell Island	Campbell Island	14, CAM	Blood	23/10/2013	Pit Tag 153723110 #027 Male	3
JW532	*Eudyptes filholi*	Campbell Island	Campbell Island	14, CAM	Blood	23/10/2013	09106579767 #020	3
JW533	*Eudyptes filholi*	Campbell Island	Campbell Island	14, CAM	Blood	9/11/2013	9106457331 GLS #01 Female	1
JW534, JW569	*Eudyptes filholi*	Campbell Island	Campbell Island	14, CAM	Blood	23/10/2013	Pit Tag 153717399 GLS #013 Male	3

Sample Code	Taxon	Location	Island Group	Island Code	Sample Type	Date Collected	Sample Identifier	Category
JW535	*Eudyptes filholi*	Campbell Island	Campbell Island	14, CAM	Blood	23/10/2013	9106655049 GLS #15 Male	1
JW536	*Eudyptes filholi*	Campbell Island	Campbell Island	14, CAM	Blood	9/11/2013	9106460121 GLS #36	1
JW537	*Eudyptes filholi*	Campbell Island	Campbell Island	14, CAM	Blood	24/10/2013	Pit Tag 9106258371 GLS #032 Male	3
JW538	*Eudyptes filholi*	Campbell Island	Campbell Island	14, CAM	Blood	23/10/2013	Pit Tag 167887703 GLS #41 Female	1
JW539	*Eudyptes filholi*	Campbell Island	Campbell Island	14, CAM	Blood	23/10/2013	Pit # 153746150 GLS #009 Male	1
JW540	*Eudyptes filholi*	Campbell Island	Campbell Island	14, CAM	Blood	25/10/2013	9106062183 GL 004	1
JW541	*Eudyptes filholi*	Campbell Island	Campbell Island	14, CAM	Blood	25/10/2013	153723083 GL 042	1
JW542	*Eudyptes filholi*	Campbell Island	Campbell Island	14, CAM	Blood	23/10/2013	9106409739 GLS #24 Male	3
JW543	*Eudyptes filholi*	Campbell Island	Campbell Island	14, CAM	Blood	23/10/2013	167887977 GLS #3 Male	3
JW544	*Eudyptes filholi*	Campbell Island	Campbell Island	14, CAM	Blood	23/10/2013	153746212 037	1
JW545, JW570	*Eudyptes filholi*	Campbell Island	Campbell Island	14, CAM	Blood	23/10/2013	Pit Tag 9106543979 GLS #030 Male	1
JW546	*Eudyptes filholi*	Campbell Island	Campbell Island	14, CAM	Blood	25/10/2013	167887535 GL 021	3
JW547, JW571	*Eudyptes filholi*	Campbell Island	Campbell Island	14, CAM	Blood	23/10/2013	153746128 GLS #7 Male	1
JW548	*Eudyptes filholi*	Campbell Island	Campbell Island	14, CAM	Blood	23/10/2013	Pit 9106622173 GLS #033 Male	3

Sample Code	Taxon	Location	Island Group	Island Code	Sample Type	Date Collected	Sample Identifier	Category
JW550	*Eudyptes filholi*	Campbell Island	Campbell Island	14, CAM	Blood	24/10/2013	Pit Tag 9106493931 GLS #016 Male	3
JW551	*Eudyptes filholi*	Campbell Island	Campbell Island	14, CAM	Blood	24/10/2013	9106547393 GL #017	1
JW552	*Eudyptes filholi*	Campbell Island	Campbell Island	14, CAM	Blood	10/11/2013	153717989 GLS #12 Female	3
JW553, JW572	*Eudyptes filholi*	Campbell Island	Campbell Island	14, CAM	Blood	24/10/2013	Pit Tag 9106542072 GLS #044 Male	1
JW554	*Eudyptes filholi*	Campbell Island	Campbell Island	14, CAM	Blood	25/10/2013	167888427 GL 025	3
JW555	*Eudyptes filholi*	Campbell Island	Campbell Island	14, CAM	Blood	23/10/2013	Pit 0910646527 GLS #005 Male	3
JW556	*Eudyptes filholi*	Campbell Island	Campbell Island	14, CAM	Blood	24/10/2013	Pit Tag 9106464281 GLS #039 Female	2
JW557	*Eudyptes filholi*	Campbell Island	Campbell Island	14, CAM	Blood	10/11/2013	153718228 GLS #14 Female	2
JW558, JW573	*Eudyptes filholi*	Campbell Island	Campbell Island	14, CAM	Blood	24/10/2013	Pit Tag 153740597 GLS #026 Male	1
JW559	*Eudyptes filholi*	Campbell Island	Campbell Island	14, CAM	Blood	9/11/2013	9106492034 GLS #28 Female	1
JW560	*Eudyptes filholi*	Campbell Island	Campbell Island	14, CAM	Blood	25/10/2013	153746623 GL 029	3
JW561	*Eudyptes filholi*	Campbell Island	Campbell Island	14, CAM	Blood	23/10/2013	Pit Tag 9106438343 GLS #40 Male	3
JW562	*Eudyptes filholi*	Campbell Island	Campbell Island	14, CAM	Blood	23/10/2013	153718728 GLS #23 Female	1

Sample Code	Taxon	Location	Island Group	Island Code	Sample Type	Date Collected	Sample Identifier	Category
JW563	*Eudyptes filholi*	Campbell Island	Campbell Island	14, CAM	Blood	9/11/2013	167892000 GLS #22 Female	2
JW564	*Eudyptes filholi*	Campbell Island	Campbell Island	14, CAM	Blood	24/10/2013	Pit Tag 153716942 GLS #038 Female	1
JW565	*Eudyptes filholi*	Campbell Island	Campbell Island	14, CAM	Blood	20/10/2013	PIT 153740428 GLS #002 Male	1
JW670, JW672	*Eudyptes filholi*	Haskell Bay	Auckland Islands	15, AUC	Blood	2012	1	3
JW671, JW673	*Eudyptes filholi*	Haskell Bay	Auckland Islands	15, AUC	Blood	2012	2	3
JW674	*Eudyptes filholi*	Haskell Bay	Auckland Islands	15, AUC	Blood	2012	3	2
JW675	*Eudyptes filholi*	Haskell Bay	Auckland Islands	15, AUC	Blood	2012	4	2
JW676	*Eudyptes filholi*	Haskell Bay	Auckland Islands	15, AUC	Blood	2012	5	2
JW677	*Eudyptes filholi*	Haskell Bay	Auckland Islands	15, AUC	Blood	2012	6	3
JW678	*Eudyptes filholi*	Haskell Bay	Auckland Islands	15, AUC	Blood	2012	7	3
JW679, JW700	*Eudyptes filholi*	Haskell Bay	Auckland Islands	15, AUC	Blood	2012	8	3
JW680, JW701	*Eudyptes filholi*	Haskell Bay	Auckland Islands	15, AUC	Blood	2012	9	3
JW681, JW702	*Eudyptes filholi*	Haskell Bay	Auckland Islands	15, AUC	Blood	2012	10	1
JW682, JW703	*Eudyptes filholi*	Haskell Bay	Auckland Islands	15, AUC	Blood	2012	11	1
JW683, JW704	*Eudyptes filholi*	Haskell Bay	Auckland Islands	15, AUC	Blood	2012	12	1
JW684, JW705	*Eudyptes filholi*	Haskell Bay	Auckland Islands	15, AUC	Blood	2012	13	3
JW685, JW706	*Eudyptes filholi*	Haskell Bay	Auckland Islands	15, AUC	Blood	2012	14	3
JW686, JW707	*Eudyptes filholi*	Haskell Bay	Auckland Islands	15, AUC	Blood	2012	15	1
JW688	*Eudyptes filholi*	Antipodes island	Antipodes island	16, ANT	Blood	2011	AB 21	1
JW689	*Eudyptes filholi*	Antipodes island	Antipodes island	16, ANT	Blood	2011	AB 22	1
JW690	*Eudyptes filholi*	Antipodes island	Antipodes island	16, ANT	Blood	2011	AB 23	3
JW691	*Eudyptes filholi*	Antipodes island	Antipodes island	16, ANT	Blood	2011	AB 24	3
JW692	*Eudyptes filholi*	Antipodes island	Antipodes island	16, ANT	Blood	2011	AB 26	3
JW693	*Eudyptes filholi*	Antipodes island	Antipodes island	16, ANT	Blood	2011	AB 27	1
JW694	*Eudyptes filholi*	Antipodes island	Antipodes island	16, ANT	Blood	2011	AB 25	1
JW695	*Eudyptes filholi*	Antipodes island	Antipodes island	16, ANT	Blood	2011	AB 28	1
JW696	*Eudyptes filholi*	Antipodes island	Antipodes island	16, ANT	Blood	2011	AB 29	1
JW697	*Eudyptes filholi*	Antipodes island	Antipodes island	16, ANT	Blood	2011	AB 30	1
JW698	*Eudyptes filholi*	Antipodes island	Antipodes island	16, ANT	Blood	2011	AB 31	1
JW699	*Eudyptes filholi*	Antipodes island	Antipodes island	16, ANT	Blood	2011	AB 32	1
JW742	*Eudyptes filholi*	Jardin Japonais	Crozet	10, CRZ	Blood	12/2003	JJRP1	3
JW743	*Eudyptes filholi*	Jardin Japonais	Crozet	10, CRZ	Blood	12/2003	JJRP2	3
JW744	*Eudyptes filholi*	Jardin Japonais	Crozet	10, CRZ	Blood	12/2003	JJRP3	3

Sample Code	Taxon	Location	Island Group	Island Code	Sample Type	Date Collected	Sample Identifier	Category
JW745	*Eudyptes filholi*	Jardin Japonais	Crozet	10, CRZ	Blood	12/2003	JJRP4	3
JW746	*Eudyptes filholi*	Jardin Japonais	Crozet	10, CRZ	Blood	12/2003	JJRP5	3
JW747	*Eudyptes filholi*	Jardin Japonais	Crozet	10, CRZ	Blood	12/2003	JJRP6	3
JW748	*Eudyptes filholi*	Jardin Japonais	Crozet	10, CRZ	Blood	12/2003	JJRP7	3
JW749	*Eudyptes filholi*	Jardin Japonais	Crozet	10, CRZ	Blood	12/2003	JJRP8	3
JW750	*Eudyptes filholi*	Jardin Japonais	Crozet	10, CRZ	Blood	12/2003	JJRP9	3
JW751	*Eudyptes filholi*	Jardin Japonais	Crozet	10, CRZ	Blood	12/2003	JJRP10	3
JW752	*Eudyptes filholi*	Jardin Japonais	Crozet	10, CRZ	Blood	12/2003	JJRP11	3
JW753	*Eudyptes filholi*	Jardin Japonais	Crozet	10, CRZ	Blood	12/2003	JJRP12	3
JW754	*Eudyptes filholi*	Jardin Japonais	Crozet	10, CRZ	Blood	12/2003	JJRP13	3
JW755	*Eudyptes filholi*	Jardin Japonais	Crozet	10, CRZ	Blood	12/2003	JJRP14	3
JW756	*Eudyptes filholi*	Jardin Japonais	Crozet	10, CRZ	Blood	12/2003	JJRP15	3
JW796, JW840	*Eudyptes filholi*	Cave Bay	Prince Edward Island	9, PEI	Blood	21/12/2008	1	3
JW797	*Eudyptes filholi*	Cave Bay	Prince Edward Island	9, PEI	Blood	21/12/2008	2	3
JW798	*Eudyptes filholi*	Cave Bay	Prince Edward Island	9, PEI	Blood	21/12/2008	3	3
JW799	*Eudyptes filholi*	Cave Bay	Prince Edward Island	9, PEI	Blood	21/12/2008	4	3
JW800	*Eudyptes filholi*	Cave Bay	Prince Edward Island	9, PEI	Blood	21/12/2008	5	3
JW801	*Eudyptes filholi*	Cave Bay	Prince Edward Island	9, PEI	Blood	21/12/2008	6	3
JW802	*Eudyptes filholi*	Cave Bay	Prince Edward Island	9, PEI	Blood	21/12/2008	7	3
JW803	*Eudyptes filholi*	Cave Bay	Prince Edward Island	9, PEI	Blood	21/12/2008	8	3
JW804	*Eudyptes filholi*	Cave Bay	Prince Edward Island	9, PEI	Blood	21/12/2008	9	3
JW805, JW841	*Eudyptes filholi*	Cave Bay	Prince Edward Island	9, PEI	Blood	21/12/2008	10	3
JW806	*Eudyptes filholi*	Cave Bay	Prince Edward Island	9, PEI	Blood	21/12/2008	11	3
JW807	*Eudyptes filholi*	Cave Bay	Prince Edward Island	9, PEI	Blood	21/12/2008	12	3
JW808, JW842	*Eudyptes filholi*	Cave Bay	Prince Edward Island	9, PEI	Blood	21/12/2008	13	3
JW809	*Eudyptes filholi*	Cave Bay	Prince Edward Island	9, PEI	Blood	21/12/2008	14	3
JW810	*Eudyptes filholi*	Cave Bay	Prince Edward Island	9, PEI	Blood	21/12/2008	15	3
JW811	*Eudyptes filholi*	Cave Bay	Prince Edward Island	9, PEI	Blood	21/12/2008	16	3
JW812, JW843	*Eudyptes filholi*	Cave Bay	Prince Edward Island	9, PEI	Blood	21/12/2008	17	3
JW813	*Eudyptes filholi*	Cave Bay	Prince Edward Island	9, PEI	Blood	21/12/2008	18	3
JW814	*Eudyptes filholi*	Cave Bay	Prince Edward Island	9, PEI	Blood	21/12/2008	19	3
JW815	*Eudyptes filholi*	Cave Bay	Prince Edward Island	9, PEI	Blood	21/12/2008	20	3
JW960	*Eudyptes filholi*	Base, Tripot	Marion Island	9, MAR	Feathers	21/04/2011	RPTRYP 1	3
JW961	*Eudyptes filholi*	Base, Tripot	Marion Island	9, MAR	Feathers	21/04/2011	RPTRYP 2	2
JW962	*Eudyptes filholi*	Base, Tripot	Marion Island	9, MAR	Feathers	21/04/2011	RPTRYP 3	1
JW963	*Eudyptes filholi*	Base, Tripot	Marion Island	9, MAR	Feathers	21/04/2011	RPTRYP 4	1
JW964	*Eudyptes filholi*	Ship's Cove	Marion Island	9, MAR	Feathers	16/04/2011	RPSHIP RH7	1
JW965	*Eudyptes filholi*	Ship's Cove	Marion Island	9, MAR	Feathers	16/04/2011	RPSHIP RH75	3
JW966	*Eudyptes filholi*	Ship's Cove	Marion Island	9, MAR	Feathers	16/04/2011	RPSHIP RH6	1
JW967	*Eudyptes filholi*	Sealer's Beach	Marion Island	9, MAR	Feathers	15/04/2011	RH15 1	1
JW968	*Eudyptes filholi*	Sealer's Beach	Marion Island	9, MAR	Feathers	25/04/2011	RH21 2	3

Sample Code	Taxon	Location	Island Group	Island Code	Sample Type	Date Collected	Sample Identifier	Category
JW969	*Eudyptes filholi*	Sealer's Beach	Marion Island	9, MAR	Feathers	13/04/2011	RH15 3	1
JW970	*Eudyptes filholi*	Sealer's Beach	Marion Island	9, MAR	Feathers	17/04/2011	RH9 4	1
JW1176	*Eudyptes filholi*	Sealer's Beach	Marion Island	9, MAR	Feathers	2011	RH2	1
JW1177	*Eudyptes filholi*	Sealer's Beach	Marion Island	9, MAR	Feathers	2011	RH18	1
JW1178	*Eudyptes filholi*	Sealer's Beach	Marion Island	9, MAR	Feathers	2011	RH13	1
JW971	*Eudyptes filholi*	Maccibay, Tripot	Marion Island	9, MAR	Feathers	21/04/2011	RPTRY 1	1
JW972	*Eudyptes filholi*	Maccibay, Tripot	Marion Island	9, MAR	Feathers	21/04/2011	RPTRY 2	1
JW973	*Eudyptes filholi*	Maccibay, Tripot	Marion Island	9, MAR	Feathers	21/04/2011	RPTRY 3	2
JW974	*Eudyptes filholi*	Maccibay, Tripot	Marion Island	9, MAR	Feathers	21/04/2011	RPTRY 4	1
JW975	*Eudyptes filholi*	Swartkops	Marion Island	9, MAR	Feathers	18/04/2011	RPSWAR 1	3
JW976	*Eudyptes filholi*	Swartkops	Marion Island	9, MAR	Feathers	18/04/2011	RPSWAR 2	1
JW977	*Eudyptes filholi*	Swartkops	Marion Island	9, MAR	Feathers	18/04/2011	RPSWAR 3	1
JW978	*Eudyptes filholi*	Swartkops	Marion Island	9, MAR	Feathers	18/04/2011	RPSWAR 4	1
JW979	*Eudyptes filholi*	Swartkops	Marion Island	9, MAR	Feathers	18/04/2011	RPSWAR 5	1
JW1193	*Eudyptes filholi*	North Precarious Bay	Macquarie Island	13, MAC	Tissue	03/2015	BBA 15 1266	3
JW1194	*Eudyptes filholi*	North Precarious Bay	Macquarie Island	13, MAC	Tissue	03/2015	BBA 1224	3
JW1195	*Eudyptes filholi*	Windsor Bay	Macquarie Island	13, MAC	Tissue	03/2015	D14 480	3
JW1197	*Eudyptes filholi*	Windsor Bay	Macquarie Island	13, MAC	Tissue	03/2015	D14 490	3
JW1196	*Eudyptes filholi*	Caroline Point	Macquarie Island	13, MAC	Tissue	03/2015	D14 492	3
JW1199	*Eudyptes filholi*	Caroline Point	Macquarie Island	13, MAC	Tissue	03/2015	D14 486	3
JW1200	*Eudyptes filholi*	Caroline Point	Macquarie Island	13, MAC	Tissue	03/2015	D14 484	3
JW1198	*Eudyptes filholi*	Hurd Point	Macquarie Island	13, MAC	Tissue	03/2015	D14 460	3
JW984, JW1043, JW1100	*Eudyptes moseleyi*	Amsterdam Island	Amsterdam Island	12, AMS	Blood	11/2011	1	1
JW985, JW1044	*Eudyptes moseleyi*	Amsterdam Island	Amsterdam Island	12, AMS	Blood	11/2011	2	3
JW986, JW1045, JW1101	*Eudyptes moseleyi*	Amsterdam Island	Amsterdam Island	12, AMS	Blood	11/2011	3	3
JW987, JW1046, JW1102	*Eudyptes moseleyi*	Amsterdam Island	Amsterdam Island	12, AMS	Blood	11/2011	4	3
JW988, JW1047	*Eudyptes moseleyi*	Amsterdam Island	Amsterdam Island	12, AMS	Blood	11/2011	5	1
JW989, JW1048	*Eudyptes moseleyi*	Amsterdam Island	Amsterdam Island	12, AMS	Blood	11/2011	6	3
JW990, JW1049, JW1103	*Eudyptes moseleyi*	Amsterdam Island	Amsterdam Island	12, AMS	Blood	11/2011	7	3
JW991, JW1050	*Eudyptes moseleyi*	Amsterdam Island	Amsterdam Island	12, AMS	Blood	11/2011	8	1
JW992, JW1051	*Eudyptes moseleyi*	Amsterdam Island	Amsterdam Island	12, AMS	Blood	11/2011	9	1
JW993, JW1052	*Eudyptes moseleyi*	Amsterdam Island	Amsterdam Island	12, AMS	Blood	11/2011	10	1
JW994, JW1053	*Eudyptes moseleyi*	Amsterdam Island	Amsterdam Island	12, AMS	Blood	11/2011	11	1
JW995, JW1054, JW1104	*Eudyptes moseleyi*	Amsterdam Island	Amsterdam Island	12, AMS	Blood	11/2011	12	1
JW996, JW1055, JW1105	*Eudyptes moseleyi*	Amsterdam Island	Amsterdam Island	12, AMS	Blood	11/2011	13	1
JW997, JW1056	*Eudyptes moseleyi*	Amsterdam Island	Amsterdam Island	12, AMS	Blood	11/2011	14	1
JW998, JW1057, JW1106	*Eudyptes moseleyi*	Amsterdam Island	Amsterdam Island	12, AMS	Blood	11/2011	15	3
JW999, JW1058	*Eudyptes moseleyi*	Amsterdam Island	Amsterdam Island	12, AMS	Blood	11/2011	16	1
JW1000, JW1059	*Eudyptes moseleyi*	Amsterdam Island	Amsterdam Island	12, AMS	Blood	11/2011	17	1
JW1001, JW1060, JW1107	*Eudyptes moseleyi*	Amsterdam Island	Amsterdam Island	12, AMS	Blood	11/2011	18	1
JW1002, JW1061	*Eudyptes moseleyi*	Amsterdam Island	Amsterdam Island	12, AMS	Blood	11/2011	19	1

Sample Code	Taxon	Location	Island Group	Island Code	Sample Type	Date Collected	Sample Identifier	Category
JW1003, JW1062	*Eudyptes moseleyi*	Amsterdam Island	Amsterdam Island	12, AMS	Blood	11/2011	20	1
JW1121	*Eudyptes moseleyi*	Admirals	Gough Island	8, GOU	Feathers	10/2010	1	1
JW1122	*Eudyptes moseleyi*	Admirals	Gough Island	8, GOU	Feathers	10/2010	2	1
JW1123	*Eudyptes moseleyi*	Admirals	Gough Island	8, GOU	Feathers	10/2010	3	3
JW1124	*Eudyptes moseleyi*	Admirals	Gough Island	8, GOU	Feathers	10/2010	4	3
JW1125	*Eudyptes moseleyi*	Admirals	Gough Island	8, GOU	Feathers	10/2010	5	1
JW1183	*Eudyptes moseleyi*	Admirals	Gough Island	8, GOU	Feathers	10/2010	6	1
JW1184	*Eudyptes moseleyi*	Admirals	Gough Island	8, GOU	Feathers	10/2010	7	1
JW758	*Eudyptes chrysolophus chrysolophus*	Jardin Japonais	Crozet	10, CRZ	Blood	12/2003	MPJJ 1	3
JW759, JW779	*Eudyptes chrysolophus chrysolophus*	Jardin Japonais	Crozet	10, CRZ	Blood	12/2003	MPJJ 2	1
JW760	*Eudyptes chrysolophus chrysolophus*	Jardin Japonais	Crozet	10, CRZ	Blood	12/2003	MPJJ 3	3
JW761	*Eudyptes chrysolophus chrysolophus*	Jardin Japonais	Crozet	10, CRZ	Blood	12/2003	MPJJ 4	3
JW762	*Eudyptes chrysolophus chrysolophus*	Jardin Japonais	Crozet	10, CRZ	Blood	12/2003	MPJJ 5	1
JW763, JW780	*Eudyptes chrysolophus chrysolophus*	Jardin Japonais	Crozet	10, CRZ	Blood	12/2003	MPJJ 6	3
JW764	*Eudyptes chrysolophus chrysolophus*	Jardin Japonais	Crozet	10, CRZ	Blood	12/2003	MPJJ 7	3
JW765	*Eudyptes chrysolophus chrysolophus*	Jardin Japonais	Crozet	10, CRZ	Blood	12/2003	MPJJ 8	3
JW766	*Eudyptes chrysolophus chrysolophus*	Jardin Japonais	Crozet	10, CRZ	Blood	12/2003	MPJJ 9	3
JW767	*Eudyptes chrysolophus chrysolophus*	Jardin Japonais	Crozet	10, CRZ	Blood	12/2003	MPJJ 10	3
JW768	*Eudyptes chrysolophus chrysolophus*	Jardin Japonais	Crozet	10, CRZ	Blood	12/2003	MPJJ 11	3
JW769, JW778	*Eudyptes chrysolophus chrysolophus*	Jardin Japonais	Crozet	10, CRZ	Blood	12/2003	MPJJ 12	1
JW770, JW781	*Eudyptes chrysolophus chrysolophus*	Jardin Japonais	Crozet	10, CRZ	Blood	12/2003	MPJJ 13	3
JW771	*Eudyptes chrysolophus chrysolophus*	Jardin Japonais	Crozet	10, CRZ	Blood	12/2003	MPJJ 14	3
JW772	*Eudyptes chrysolophus chrysolophus*	Jardin Japonais	Crozet	10, CRZ	Blood	12/2003	MPJJ 15	3
JW773	*Eudyptes chrysolophus chrysolophus*	Jardin Japonais	Crozet	10, CRZ	Blood	12/2003	MPJJ 16	3

Sample Code	Taxon	Location	Island Group	Island Code	Sample Type	Date Collected	Sample Identifier	Category
JW774	*Eudyptes chrysolophus chrysolophus*	Jardin Japonais	Crozet	10, CRZ	Blood	12/2003	MPJJ 17	3
JW775	*Eudyptes chrysolophus chrysolophus*	Jardin Japonais	Crozet	10, CRZ	Blood	12/2003	MPJJ 18	3
JW776	*Eudyptes chrysolophus chrysolophus*	Jardin Japonais	Crozet	10, CRZ	Blood	12/2003	MPJJ 19	3
JW777	*Eudyptes chrysolophus chrysolophus*	Jardin Japonais	Crozet	10, CRZ	Blood	12/2003	MPJJ 20	3
JW820	*Eudyptes chrysolophus chrysolophus*	RSA Point	Prince Edward Island	9, PEI	Blood	12/2008	1	3
JW821	*Eudyptes chrysolophus chrysolophus*	RSA Point	Prince Edward Island	9, PEI	Blood	12/2008	2	3
JW822, JW864	*Eudyptes chrysolophus chrysolophus*	RSA Point	Prince Edward Island	9, PEI	Blood	12/2008	3	2
JW823	*Eudyptes chrysolophus chrysolophus*	RSA Point	Prince Edward Island	9, PEI	Blood	12/2008	4	3
JW824, JW865	*Eudyptes chrysolophus chrysolophus*	RSA Point	Prince Edward Island	9, PEI	Blood	12/2008	5	3
JW825	*Eudyptes chrysolophus chrysolophus*	RSA Point	Prince Edward Island	9, PEI	Blood	12/2008	6	3
JW826	*Eudyptes chrysolophus chrysolophus*	RSA Point	Prince Edward Island	9, PEI	Blood	12/2008	7	3
JW827	*Eudyptes chrysolophus chrysolophus*	RSA Point	Prince Edward Island	9, PEI	Blood	12/2008	8	3
JW828, JW866	*Eudyptes chrysolophus chrysolophus*	RSA Point	Prince Edward Island	9, PEI	Blood	12/2008	9	3
JW829, JW867	*Eudyptes chrysolophus chrysolophus*	RSA Point	Prince Edward Island	9, PEI	Blood	12/2008	10	3
JW830	*Eudyptes chrysolophus chrysolophus*	RSA Point	Prince Edward Island	9, PEI	Blood	12/2008	11	3
JW831	*Eudyptes chrysolophus chrysolophus*	RSA Point	Prince Edward Island	9, PEI	Blood	12/2008	12	3
JW832	*Eudyptes chrysolophus chrysolophus*	RSA Point	Prince Edward Island	9, PEI	Blood	12/2008	13	3
JW833, JW863	*Eudyptes chrysolophus chrysolophus*	RSA Point	Prince Edward Island	9, PEI	Blood	12/2008	14	3
JW834	*Eudyptes chrysolophus chrysolophus*	RSA Point	Prince Edward Island	9, PEI	Blood	12/2008	15	3
JW835	*Eudyptes chrysolophus chrysolophus*	RSA Point	Prince Edward Island	9, PEI	Blood	12/2008	16	3

Sample Code	Taxon	Location	Island Group	Island Code	Sample Type	Date Collected	Sample Identifier	Category
JW836	*Eudyptes chrysolophus chrysolophus*	RSA Point	Prince Edward Island	9, PEI	Blood	12/2008	17	3
JW837	*Eudyptes chrysolophus chrysolophus*	RSA Point	Prince Edward Island	9, PEI	Blood	12/2008	18	3
JW838, JW888	*Eudyptes chrysolophus chrysolophus*	RSA Point	Prince Edward Island	9, PEI	Blood	12/2008	19	3
JW839	*Eudyptes chrysolophus chrysolophus*	RSA Point	Prince Edward Island	9, PEI	Blood	12/2008	20	3
JW844	*Eudyptes chrysolophus chrysolophus*	Cooper Bay	South Georgia	5, SGE	Blood	2016	1	3
JW845	*Eudyptes chrysolophus chrysolophus*	Cooper Bay	South Georgia	5, SGE	Blood	2016	2	1
JW846	*Eudyptes chrysolophus chrysolophus*	Cooper Bay	South Georgia	5, SGE	Blood	2016	3	3
JW847	*Eudyptes chrysolophus chrysolophus*	Cooper Bay	South Georgia	5, SGE	Blood	2016	4	3
JW848	*Eudyptes chrysolophus chrysolophus*	Cooper Bay	South Georgia	5, SGE	Blood	2016	5	1
JW849	*Eudyptes chrysolophus chrysolophus*	Cooper Bay	South Georgia	5, SGE	Blood	2016	6	3
JW850	*Eudyptes chrysolophus chrysolophus*	Cooper Bay	South Georgia	5, SGE	Blood	2016	7	3
JW851	*Eudyptes chrysolophus chrysolophus*	Cooper Bay	South Georgia	5, SGE	Blood	2016	8	1
JW852	*Eudyptes chrysolophus chrysolophus*	Cooper Bay	South Georgia	5, SGE	Blood	2016	9	3
JW853	*Eudyptes chrysolophus chrysolophus*	Cooper Bay	South Georgia	5, SGE	Blood	2016	11	3
JW854	*Eudyptes chrysolophus chrysolophus*	Cooper Bay	South Georgia	5, SGE	Blood	2016	12	1
JW855	*Eudyptes chrysolophus chrysolophus*	Cooper Bay	South Georgia	5, SGE	Blood	2016	13	3
JW856	*Eudyptes chrysolophus chrysolophus*	Cooper Bay	South Georgia	5, SGE	Blood	2016	14	3
JW857	*Eudyptes chrysolophus chrysolophus*	Cooper Bay	South Georgia	5, SGE	Blood	2016	15	1
JW858	*Eudyptes chrysolophus chrysolophus*	Cooper Bay	South Georgia	5, SGE	Blood	2016	16	1
JW859	*Eudyptes chrysolophus chrysolophus*	Cooper Bay	South Georgia	5, SGE	Blood	2016	17	3

Sample Code	Taxon	Location	Island Group	Island Code	Sample Type	Date Collected	Sample Identifier	Category
JW860	*Eudyptes chrysolophus chrysolophus*	Cooper Bay	South Georgia	5, SGE	Blood	2016	18	3
JW861	*Eudyptes chrysolophus chrysolophus*	Cooper Bay	South Georgia	5, SGE	Blood	2016	19	3
JW862	*Eudyptes chrysolophus chrysolophus*	Cooper Bay	South Georgia	5, SGE	Blood	2016	20	3
JW928	*Eudyptes chrysolophus chrysolophus*	Zavodovski	South Sandwich Islands	6, SSI	Feathers	01/2011	MPZ1 014	3
JW929	*Eudyptes chrysolophus chrysolophus*	Zavodovski	South Sandwich Islands	6, SSI	Feathers	01/2011	MPZ1 005	1
JW930	*Eudyptes chrysolophus chrysolophus*	Zavodovski	South Sandwich Islands	6, SSI	Feathers	01/2011	MPZ1 003	1
JW931	*Eudyptes chrysolophus chrysolophus*	Zavodovski	South Sandwich Islands	6, SSI	Feathers	01/2011	MPZ1 006	3
JW932	*Eudyptes chrysolophus chrysolophus*	Zavodovski	South Sandwich Islands	6, SSI	Feathers	01/2011	No Label	3
JW1156	*Eudyptes chrysolophus chrysolophus*	Zavodovski	South Sandwich Islands	6, SSI	Feathers	01/2011	MPZ1 024	1
JW1157	*Eudyptes chrysolophus chrysolophus*	Zavodovski	South Sandwich Islands	6, SSI	Feathers	01/2011	MPZ1 022	1
JW1158	*Eudyptes chrysolophus chrysolophus*	Zavodovski	South Sandwich Islands	6, SSI	Feathers	01/2011	MPZ1 028	1
JW1159	*Eudyptes chrysolophus chrysolophus*	Zavodovski	South Sandwich Islands	6, SSI	Feathers	01/2011	MPZ1 034	1
JW1160	*Eudyptes chrysolophus chrysolophus*	Zavodovski	South Sandwich Islands	6, SSI	Feathers	01/2011	MPZ1 009	1
JW1161	*Eudyptes chrysolophus chrysolophus*	Zavodovski	South Sandwich Islands	6, SSI	Feathers	01/2011	MPZ1 038	1
JW938	*Eudyptes chrysolophus chrysolophus*	Clapmach Point, Candlemas	South Sandwich Islands	6, SSI	Feathers	01/2011	MPCMP 018	1
JW939	*Eudyptes chrysolophus chrysolophus*	Clapmach Point, Candlemas	South Sandwich Islands	6, SSI	Feathers	01/2011	MPCMP 033	1
JW940	*Eudyptes chrysolophus chrysolophus*	Clapmach Point, Candlemas	South Sandwich Islands	6, SSI	Feathers	01/2011	MPCMP 008	1
JW941	*Eudyptes chrysolophus chrysolophus*	Clapmach Point, Candlemas	South Sandwich Islands	6, SSI	Feathers	01/2011	MPCMP 013	1
JW942	*Eudyptes chrysolophus chrysolophus*	Clapmach Point, Candlemas	South Sandwich Islands	6, SSI	Feathers	01/2011	MPCMP 014	3
JW1164	*Eudyptes chrysolophus chrysolophus*	Clapmach Point, Candlemas	South Sandwich Islands	6, SSI	Feathers	01/2011	MP CMP 002	1

Sample Code	Taxon	Location	Island Group	Island Code	Sample Type	Date Collected	Sample Identifier	Category
JW1165	*Eudyptes chrysolophus chrysolophus*	Clapmach Point, Candlemas	South Sandwich Islands	6, SSI	Feathers	01/2011	MP CMP 041	1
JW1166	*Eudyptes chrysolophus chrysolophus*	Clapmach Point, Candlemas	South Sandwich Islands	6, SSI	Feathers	01/2011	MP CMP 004	1
JW1167	*Eudyptes chrysolophus chrysolophus*	Clapmach Point, Candlemas	South Sandwich Islands	6, SSI	Feathers	01/2011	MP CMP 001	1
JW1168	*Eudyptes chrysolophus chrysolophus*	Clapmach Point, Candlemas	South Sandwich Islands	6, SSI	Feathers	01/2011	MP CMP 007	1
JW944	*Eudyptes chrysolophus chrysolophus*	Swartkops	Marion Island	9, MAR	Feathers	18/04/2011	MPSWAR 1	3
JW945	*Eudyptes chrysolophus chrysolophus*	Swartkops	Marion Island	9, MAR	Feathers	18/04/2011	MPSWAR 2	3
JW946	*Eudyptes chrysolophus chrysolophus*	Swartkops	Marion Island	9, MAR	Feathers	18/04/2011	MPSWAR 3	1
JW947	*Eudyptes chrysolophus chrysolophus*	Swartkops	Marion Island	9, MAR	Feathers	18/04/2011	MPSWAR 4	3
JW948	*Eudyptes chrysolophus chrysolophus*	Swartkops	Marion Island	9, MAR	Feathers	18/04/2011	MPSWAR 5	3
JW949	*Eudyptes chrysolophus chrysolophus*	Ship's Cove	Marion Island	9, MAR	Feathers	16/04/2011	MPSHIP Mac 30	3
JW950	*Eudyptes chrysolophus chrysolophus*	Ship's Cove	Marion Island	9, MAR	Feathers	16/04/2011	MPSHIP Mac 26	3
JW951	*Eudyptes chrysolophus chrysolophus*	Ship's Cove	Marion Island	9, MAR	Feathers	16/04/2011	MPSHIP Mac 27	3
JW952	*Eudyptes chrysolophus chrysolophus*	Ship's Cove	Marion Island	9, MAR	Feathers	16/04/2011	MPSHIP Mac 19	1
JW953	*Eudyptes chrysolophus chrysolophus*	Ship's Cove	Marion Island	9, MAR	Feathers	16/04/2011	MPSHIP Mac 22	3
JW954	*Eudyptes chrysolophus chrysolophus*	Bullard Beach	Marion Island	9, MAR	Feathers	13/04/2011	MPBULL Maci 9	1
JW955	*Eudyptes chrysolophus chrysolophus*	Bullard Beach	Marion Island	9, MAR	Feathers	13/04/2011	MPBULL Maci 6	3
JW956	*Eudyptes chrysolophus chrysolophus*	Bullard Beach	Marion Island	9, MAR	Feathers	13/04/2011	MPBULL Maci 3	3
JW957	*Eudyptes chrysolophus chrysolophus*	Bullard Beach	Marion Island	9, MAR	Feathers	13/04/2011	MPBULL Maci 1	1
JW958	*Eudyptes chrysolophus chrysolophus*	Bullard Beach	Marion Island	9, MAR	Feathers	13/04/2011	MPBULL Maci 2	3
JW1004, JW1063, JW1108	*Eudyptes chrysolophus chrysolophus*	Kerguelen	Kerguelen	11, KER	Blood	01/2014	1	2

Sample Code	Taxon	Location	Island Group	Island Code	Sample Type	Date Collected	Sample Identifier	Category
JW1025, JW1079	*Eudyptes chrysolophus chrysolophus*	Kerguelen	Kerguelen	11, KER	Blood	01/2014	22	1
JW1026	*Eudyptes chrysolophus chrysolophus*	Seal Bay	Elephant Island	3, ELE	Blood	12/2015	SEAL128	3
JW1027	*Eudyptes chrysolophus chrysolophus*	Seal Bay	Elephant Island	3, ELE	Blood	12/2015	SEAL129	3
JW1028	*Eudyptes chrysolophus chrysolophus*	Seal Bay	Elephant Island	3, ELE	Blood	12/2015	SEAL130	3
JW1029	*Eudyptes chrysolophus chrysolophus*	Seal Bay	Elephant Island	3, ELE	Blood	12/2015	SEAL131	1
JW1030	*Eudyptes chrysolophus chrysolophus*	Seal Bay	Elephant Island	3, ELE	Blood	12/2015	SEAL132	3
JW1031	*Eudyptes chrysolophus chrysolophus*	Seal Bay	Elephant Island	3, ELE	Blood	12/2015	SEAL133	3
JW1032	*Eudyptes chrysolophus chrysolophus*	Seal Bay	Elephant Island	3, ELE	Blood	12/2015	SEAL134	3
JW1033	*Eudyptes chrysolophus chrysolophus*	Seal Bay	Elephant Island	3, ELE	Blood	12/2015	SEAL135	3
JW1034	*Eudyptes chrysolophus chrysolophus*	Seal Bay	Elephant Island	3, ELE	Blood	12/2015	SEAL136	3
JW1035	*Eudyptes chrysolophus chrysolophus*	Seal Bay	Elephant Island	3, ELE	Blood	12/2015	SEAL137	3
JW1036	*Eudyptes chrysolophus chrysolophus*	Seal Bay	Elephant Island	3, ELE	Blood	12/2015	SEAL138	3
JW1037	*Eudyptes chrysolophus chrysolophus*	Seal Bay	Elephant Island	3, ELE	Blood	12/2015	SEAL139	3
JW1038	*Eudyptes chrysolophus chrysolophus*	Seal Bay	Elephant Island	3, ELE	Blood	12/2015	SEAL140	3
JW1039	*Eudyptes chrysolophus chrysolophus*	Seal Bay	Elephant Island	3, ELE	Blood	12/2015	SEAL141	3
JW1040	*Eudyptes chrysolophus chrysolophus*	Seal Bay	Elephant Island	3, ELE	Blood	12/2015	SEAL142	3
JW1041	*Eudyptes chrysolophus chrysolophus*	Seal Bay	Elephant Island	3, ELE	Blood	12/2015	SEAL143	1
JW1042	*Eudyptes chrysolophus chrysolophus*	Seal Bay	Elephant Island	3, ELE	Blood	12/2015	SEAL144	3
JW868	*Eudyptes chrysolophus schlegeli*	Green Gorge	Macquarie Island	13, MAC	Blood	2006	ROPE4458	3
JW869	*Eudyptes chrysolophus schlegeli*	Green Gorge	Macquarie Island	13, MAC	Blood	2006	ROPE4461	3

Sample Code	Taxon	Location	Island Group	Island Code	Sample Type	Date Collected	Sample Identifier	Category
JW1005, JW1064, JW1109	*Eudyptes chrysolophus chrysolophus*	Kerguelen	Kerguelen	11, KER	Blood	01/2014	2	1
JW1006	*Eudyptes chrysolophus chrysolophus*	Kerguelen	Kerguelen	11, KER	Blood	01/2014	3	3
JW1007, JW1065	*Eudyptes chrysolophus chrysolophus*	Kerguelen	Kerguelen	11, KER	Blood	01/2014	4	3
JW1008, JW1066, JW1110	*Eudyptes chrysolophus chrysolophus*	Kerguelen	Kerguelen	11, KER	Blood	01/2014	5	2
JW1009, JW1067, JW1111	*Eudyptes chrysolophus chrysolophus*	Kerguelen	Kerguelen	11, KER	Blood	01/2014	6	1
JW1010, JW1068, JW1112	*Eudyptes chrysolophus chrysolophus*	Kerguelen	Kerguelen	11, KER	Blood	01/2014	7	2
JW1011	*Eudyptes chrysolophus chrysolophus*	Kerguelen	Kerguelen	11, KER	Blood	01/2014	8	3
JW1012, JW1069	*Eudyptes chrysolophus chrysolophus*	Kerguelen	Kerguelen	11, KER	Blood	01/2014	9	1
JW1013, JW1070, JW1113	*Eudyptes chrysolophus chrysolophus*	Kerguelen	Kerguelen	11, KER	Blood	01/2014	10	2
JW1014, JW1071	*Eudyptes chrysolophus chrysolophus*	Kerguelen	Kerguelen	11, KER	Blood	01/2014	11	1
JW1015, JW1072	*Eudyptes chrysolophus chrysolophus*	Kerguelen	Kerguelen	11, KER	Blood	01/2014	12	2
JW1016, JW1073	*Eudyptes chrysolophus chrysolophus*	Kerguelen	Kerguelen	11, KER	Blood	01/2014	13	3
JW1017, JW1074	*Eudyptes chrysolophus chrysolophus*	Kerguelen	Kerguelen	11, KER	Blood	01/2014	14	1
JW1018, JW1075	*Eudyptes chrysolophus chrysolophus*	Kerguelen	Kerguelen	11, KER	Blood	01/2014	15	1
JW1019	*Eudyptes chrysolophus chrysolophus*	Kerguelen	Kerguelen	11, KER	Blood	01/2014	16	3
JW1020	*Eudyptes chrysolophus chrysolophus*	Kerguelen	Kerguelen	11, KER	Blood	01/2014	17	3
JW1021, JW1076, JW1114	*Eudyptes chrysolophus chrysolophus*	Kerguelen	Kerguelen	11, KER	Blood	01/2014	18	1
JW1022	*Eudyptes chrysolophus chrysolophus*	Kerguelen	Kerguelen	11, KER	Blood	01/2014	19	3
JW1023, JW1077, JW1115	*Eudyptes chrysolophus chrysolophus*	Kerguelen	Kerguelen	11, KER	Blood	01/2014	20	2
JW1024, JW1078	*Eudyptes chrysolophus chrysolophus*	Kerguelen	Kerguelen	11, KER	Blood	01/2014	21	1

Sample Code	Taxon	Location	Island Group	Island Code	Sample Type	Date Collected	Sample Identifier	Category
JW870	*Eudyptes chrysolophus schlegeli*	Green Gorge	Macquarie Island	13, MAC	Blood	2006	ROPE4449	3
JW871	*Eudyptes chrysolophus schlegeli*	Green Gorge	Macquarie Island	13, MAC	Blood	2006	ROPE4456	3
JW872	*Eudyptes chrysolophus schlegeli*	Green Gorge	Macquarie Island	13, MAC	Blood	2006	ROPE4433	3
JW873	*Eudyptes chrysolophus schlegeli*	Green Gorge	Macquarie Island	13, MAC	Blood	2006	ROPE4450	1
JW874	*Eudyptes chrysolophus schlegeli*	Green Gorge	Macquarie Island	13, MAC	Blood	2006	ROPE4441	3
JW875	*Eudyptes chrysolophus schlegeli*	Green Gorge	Macquarie Island	13, MAC	Blood	2006	ROPE4465	3
JW876	*Eudyptes chrysolophus schlegeli*	Green Gorge	Macquarie Island	13, MAC	Blood	2006	ROPE4453	3
JW877	*Eudyptes chrysolophus schlegeli*	Green Gorge	Macquarie Island	13, MAC	Blood	2006	ROPE4437	3
JW878	*Eudyptes chrysolophus schlegeli*	Green Gorge	Macquarie Island	13, MAC	Blood	2006	ROPE4431	3
JW879	*Eudyptes chrysolophus schlegeli*	Green Gorge	Macquarie Island	13, MAC	Blood	2006	ROPE4443	3
JW880	*Eudyptes chrysolophus schlegeli*	Green Gorge	Macquarie Island	13, MAC	Blood	2006	ROPE4444	3
JW881	*Eudyptes chrysolophus schlegeli*	Green Gorge	Macquarie Island	13, MAC	Blood	2006	ROPE4470	1
JW882	*Eudyptes chrysolophus schlegeli*	Green Gorge	Macquarie Island	13, MAC	Blood	2006	ROPE4436	3
JW883	*Eudyptes chrysolophus schlegeli*	Green Gorge	Macquarie Island	13, MAC	Blood	2006	ROPE4447	3
JW884	*Eudyptes chrysolophus schlegeli*	Green Gorge	Macquarie Island	13, MAC	Blood	2006	ROPE4466	3
JW885	*Eudyptes chrysolophus schlegeli*	Green Gorge	Macquarie Island	13, MAC	Blood	2006	ROPE4455	3
JW886	*Eudyptes chrysolophus schlegeli*	Green Gorge	Macquarie Island	13, MAC	Blood	2006	ROPE4445	1
JW887, JW909	*Eudyptes chrysolophus schlegeli*	Green Gorge	Macquarie Island	13, MAC	Blood	2006	ROPE4432	1
JW889	*Eudyptes chrysolophus schlegeli*	Nuggets Colony	Macquarie Island	13, MAC	Blood	2006	ROPE4367a	3
JW890	*Eudyptes chrysolophus schlegeli*	Nuggets Colony	Macquarie Island	13, MAC	Blood	2006	ROPE4346	3

Sample Code	Taxon	Location	Island Group	Island Code	Sample Type	Date Collected	Sample Identifier	Category
JW891	*Eudyptes chrysolophus schlegeli*	Nuggets Colony	Macquarie Island	13, MAC	Blood	2006	ROPE4350	2
JW892	*Eudyptes chrysolophus schlegeli*	Nuggets Colony	Macquarie Island	13, MAC	Blood	2006	ROPE4344	3
JW893	*Eudyptes chrysolophus schlegeli*	Nuggets Colony	Macquarie Island	13, MAC	Blood	2006	ROPE4368	3
JW894	*Eudyptes chrysolophus schlegeli*	Nuggets Colony	Macquarie Island	13, MAC	Blood	2006	ROPE4370	3
JW895	*Eudyptes chrysolophus schlegeli*	Nuggets Colony	Macquarie Island	13, MAC	Blood	2006	ROPE4374	3
JW896	*Eudyptes chrysolophus schlegeli*	Nuggets Colony	Macquarie Island	13, MAC	Blood	2006	ROPE4757	3
JW897	*Eudyptes chrysolophus schlegeli*	Nuggets Colony	Macquarie Island	13, MAC	Blood	2006	ROPE4356	3
JW898	*Eudyptes chrysolophus schlegeli*	Nuggets Colony	Macquarie Island	13, MAC	Blood	2006	ROPE4341	3
JW899	*Eudyptes chrysolophus schlegeli*	Nuggets Colony	Macquarie Island	13, MAC	Blood	2006	ROPE4360	3
JW900	*Eudyptes chrysolophus schlegeli*	Nuggets Colony	Macquarie Island	13, MAC	Blood	2006	ROPE4367b	3
JW901	*Eudyptes chrysolophus schlegeli*	Nuggets Colony	Macquarie Island	13, MAC	Blood	2006	ROPE4353	3
JW902	*Eudyptes chrysolophus schlegeli*	Nuggets Colony	Macquarie Island	13, MAC	Blood	2006	ROPE4339	3
JW903	*Eudyptes chrysolophus schlegeli*	Nuggets Colony	Macquarie Island	13, MAC	Blood	2006	ROPE4369	3
JW904	*Eudyptes chrysolophus schlegeli*	Nuggets Colony	Macquarie Island	13, MAC	Blood	2006	ROPE4329	3
JW905	*Eudyptes chrysolophus schlegeli*	Nuggets Colony	Macquarie Island	13, MAC	Blood	2006	ROPE4355	3
JW906, JW927	*Eudyptes chrysolophus schlegeli*	Nuggets Colony	Macquarie Island	13, MAC	Blood	2006	ROPE4358	3
JW907	*Eudyptes chrysolophus schlegeli*	Nuggets Colony	Macquarie Island	13, MAC	Blood	2006	ROPE4349	3
JW908	*Eudyptes chrysolophus schlegeli*	Nuggets Colony	Macquarie Island	13, MAC	Blood	2006	ROPE4328	3
JW782, JW783, JW784	*Eudyptes chrysolophus* ssp. 'white faced'	Kildalkey	Marion Island	9, MAR	Blood	23/12/2008	1	3
JW785, JW786	*Eudyptes chrysolophus* ssp. 'white faced'	Kildalkey	Marion Island	9, MAR	Blood	23/12/2008	2	3

Sample Code	Taxon	Location	Island Group	Island Code	Sample Type	Date Collected	Sample Identifier	Category
JW787, JW788, JW789	*Eudyptes chrysolophus* ssp. 'white faced'	Kildalkey	Marion Island	9, MAR	Blood	23/12/2008	3	1
JW790, JW791	*Eudyptes chrysolophus* ssp. 'white faced'	Kildalkey	Marion Island	9, MAR	Blood	23/12/2008	4	1
JW792, JW793	*Eudyptes chrysolophus* ssp. 'white faced'	Kildalkey	Marion Island	9, MAR	Blood	23/12/2008	5	3
JW794, JW795	*Eudyptes chrysolophus* ssp. 'white faced'	Funk Bay	Marion Island	9, MAR	Blood	23/12/2008	1	3
JW310, JW420	*Eudyptes pachyrhynchus*	Codfish Island	Codfish Island	18, COD	Blood	22/09/2016	WH01 2016	1
JW311, JW421, JW638	*Eudyptes pachyrhynchus*	Codfish Island	Codfish Island	18, COD	Blood	29/09/2016	WH02 2016	3
JW312, JW422, JW639	*Eudyptes pachyrhynchus*	Codfish Island	Codfish Island	18, COD	Blood	22/09/2016	WH11 2016	3
JW313, JW640	*Eudyptes pachyrhynchus*	Codfish Island	Codfish Island	18, COD	Blood	29/09/2016	WH11A 2016	3
JW314, JW641	*Eudyptes pachyrhynchus*	Codfish Island	Codfish Island	18, COD	Blood	29/09/2016	WH13 2016	3
JW315, JW423, JW642, JW910	*Eudyptes pachyrhynchus*	Codfish Island	Codfish Island	18, COD	Blood	23/09/2016	WH16 2016	1
JW316, JW643	*Eudyptes pachyrhynchus*	Codfish Island	Codfish Island	18, COD	Blood	29/09/2016	WH17 2016	3
JW317, JW424, JW644, JW911	*Eudyptes pachyrhynchus*	Codfish Island	Codfish Island	18, COD	Blood	30/09/2016	WH18 2016	1
JW318, JW425, JW645	*Eudyptes pachyrhynchus*	Codfish Island	Codfish Island	18, COD	Blood	24/09/2016	WH29 2016	3
JW319, JW426, JW646, JW912	*Eudyptes pachyrhynchus*	Codfish Island	Codfish Island	18, COD	Blood	23/09/2016	WH31 2016	3
JW320, JW647	*Eudyptes pachyrhynchus*	Codfish Island	Codfish Island	18, COD	Blood	29/09/2016	WH33 2016	3
JW322, JW427, JW648, JW913	*Eudyptes pachyrhynchus*	Codfish Island	Codfish Island	18, COD	Blood	22/09/2016	WH12 2016	3
JW323, JW428, JW649, JW914	*Eudyptes pachyrhynchus*	Codfish Island	Codfish Island	18, COD	Blood	28/09/2016	WH27 2016	1
JW324, JW429, JW650, JW915	*Eudyptes pachyrhynchus*	Codfish Island	Codfish Island	18, COD	Blood	23/09/2016	WH28 2016	3
JW325, JW430, JW651	*Eudyptes pachyrhynchus*	Jackson Head, West Coast	New Zealand	18, JAC	Blood	2016	JH01 2015	1
JW326, JW652	*Eudyptes pachyrhynchus*	Jackson Head, West Coast	New Zealand	18, JAC	Blood	2016	JH01 2015 (2)	1

Sample Code	Taxon	Location	Island Group	Island Code	Sample Type	Date Collected	Sample Identifier	Category
JW327, JW431, JW653, JW917	*Eudyptes pachyrhynchus*	Jackson Head, West Coast	New Zealand	18, JAC	Blood	2016	JH 2016 Cole1	1
JW328, JW654	*Eudyptes pachyrhynchus*	Jackson Head, West Coast	New Zealand	18, JAC	Blood	2016	JH09 2014	3
JW329, JW432, JW655	*Eudyptes pachyrhynchus*	Jackson Head, West Coast	New Zealand	18, JAC	Blood	18/06/2014	JH10	3
JW330, JW433, JW656, JW918	*Eudyptes pachyrhynchus*	Jackson Head, West Coast	New Zealand	18, JAC	Blood	7/10/2016	JH21	3
JW331, JW657, JW919	*Eudyptes pachyrhynchus*	Jackson Head, West Coast	New Zealand	18, JAC	Blood	2015	JH 2015	3
JW332, JW434, JW658, JW920	*Eudyptes pachyrhynchus*	Jackson Head, West Coast	New Zealand	18, JAC	Blood	7/10/2016	JH Cave 2016	1
JW334, JW435, JW659	*Eudyptes pachyrhynchus*	Jackson Head, West Coast	New Zealand	18, JAC	Blood	2/01/2016	JH 2016	3
JW436	*Eudyptes pachyrhynchus*	Jackson Head, West Coast	New Zealand	18, JAC	Blood	8/10/2016	JH 04 2016 Female	1
JW335, JW660, JW921	*Eudyptes pachyrhynchus*	Milford Sound, Southland	New Zealand	18, MIL	Blood	2016	MS	3
JW336, JW661	*Eudyptes pachyrhynchus*	Milford Sound, Southland	New Zealand	18, MIL	Blood	2016	MS1	3
JW337, JW437, JW662, JW922	*Eudyptes pachyrhynchus*	Milford Sound, Southland	New Zealand	18, MIL	Blood	2016	MS2	1
JW338, JW438, JW663	*Eudyptes pachyrhynchus*	Milford Sound, Southland	New Zealand	18, MIL	Blood	2016	MS3	3
JW339, JW664	*Eudyptes pachyrhynchus*	Milford Sound, Southland	New Zealand	18, MIL	Blood	2016	MS4	3
JW340, JW439	*Eudyptes pachyrhynchus*	Milford Sound, Southland	New Zealand	18, MIL	Blood	2016	MS5	3
JW341, JW440, JW665, JW923	*Eudyptes pachyrhynchus*	Milford Sound, Southland	New Zealand	18, MIL	Blood	2016	MS6	2
JW342, JW666, JW924	*Eudyptes pachyrhynchus*	Milford Sound, Southland	New Zealand	18, MIL	Blood	2016	MS7	3
JW343	*Eudyptes pachyrhynchus*	Milford Sound, Southland	New Zealand	18, MIL	Blood	2016	MS8	3
JW344	*Eudyptes pachyrhynchus*	Milford Sound, Southland	New Zealand	18, MIL	Blood	2016	MS9	3
JW345, JW441, JW667, JW925	*Eudyptes pachyrhynchus*	Milford Sound, Southland	New Zealand	18, MIL	Blood	2016	MS10	2
JW346, JW442, JW668, JW926	*Eudyptes pachyrhynchus*	Milford Sound, Southland	New Zealand	18, MIL	Blood	2016	MS11	2
JW120, JW480, JW708	*Eudyptes robustus*	South Promontory	The Snares	17, SNA	Blood	11/2013	SCP 1	3

Sample Code	Taxon	Location	Island Group	Island Code	Sample Type	Date Collected	Sample Identifier	Category
JW121, JW481, JW709	*Eudyptes robustus*	South Promontory	The Snares	17, SNA	Blood	11/2013	SCP 2	1
JW122, JW482, JW710	*Eudyptes robustus*	North Promontory	The Snares	17, SNA	Blood	11/2013	SCP 3	3
JW123, JW483, JW711	*Eudyptes robustus*	North Promontory	The Snares	17, SNA	Blood	11/2013	SCP 4	1
JW124, JW712	*Eudyptes robustus*	Colony 3	The Snares	17, SNA	Blood	11/2013	SCP 5	3
JW125, JW713	*Eudyptes robustus*	Colony 3	The Snares	17, SNA	Blood	11/2013	SCP 6	3
JW126, JW484, JW714	*Eudyptes robustus*	Western Chain Toru	The Snares	17, WES	Blood	11/2013	SCP 7	2
JW127, JW485, JW715	*Eudyptes robustus*	Western Chain Toru	The Snares	17, WES	Blood	11/2013	SCP 8	3
JW128, JW486, JW716	*Eudyptes robustus*	Western Chain Toru	The Snares	17, WES	Blood	11/2013	SCP 9	1
JW129, JW717	*Eudyptes robustus*	Western Chain Toru	The Snares	17, WES	Blood	11/2013	SCP 10	1

Supplementary Table 15. Number of SNPs for 15 *Eudyptes*, *Aptenodytes* and *Pygoscelis* datasets. *n* genetic structure and genetic diversity and *n* SNPs genetic structure and genetic diversity were used to generate the Global F_{ST} and corresponding *P*-value (this table), as well as genetic diversity indices, pairwise F_{ST}, PCoA, Structure, DAPC, RAxML and SNAPP analyses (except *P. papua*, see Clucas *et al*., 2018). Global F_{ST} and the corresponding *P*-value for *Aptenodytes patagonicus*, *A. forsteri*, *Pygoscelis adeliae* and *P. antarctica* penguins are from Clucas *et al*. (2016), Younger *et al*. (2017) and Clucas *et al*. (2018). Note that some *Eudyptes* individuals were removed in the filtering process, depending on the analysis. Note also that the *P. papua* dataset has been re-filtered from Clucas *et al*. (2018) to only include samples from the South Shetland Islands and Jougla and George's Point on the West Antarctic Peninsula (see materials and methods and *Figure 13*), except for the Structure and SNAPP analysis (which were undertaken in Clucas *et al*., 2018, see also *Figure 13*), which instead included all *P. papua* individuals spanning six localities. *n* demographic and *n* SNPs demographic refers to the further filtered datasets that were used in the demographic analyses of CubSFS.

Dataset	*n* structure and genetic divesity	*n* SNPs for structure and genetic divesity	*n* demographic	*n* SNPs demographic	Global F_{ST}	*P*-value
All *Eudyptes*	254	12,988	NA	NA	0.422	0.001
Eudyptes moseleyi, *Eudyptes filholi* and *Eudyptes chrysocome*	105	10,437	NA	NA	0.154	0.001
Eudyptes moseleyi	8	4654	8	4652	0.062	0.046
Eudyptes filholi and *Eudyptes chrysocome*	97	10,694	NA	NA	0.091	0.001
Eudyptes filholi	71	8835	71	6252	0.021	0.001
Eudyptes chrysocome	26	7575	26	7573	0.006	0.021
Eudyptes chrysolophus chrysolophus and *Eudyptes chrysolophus schlegeli*	125	13,256	125	6773	0.011	0.001
Eudyptes pachyrhynchus and *Eudyptes robustus*	27	4475	NA	NA	0.163	0.001
Eudyptes pachyrhynchus	21	4279	21	4278	0.007	0.001
Eudyptes robustus	5	3141	4	3015	0.007	0.390
Aptenodytes patagonicus	64	5154	64	4627	0.002	0.001
Aptenodytes forsteri	110	4596	106	4201	0.003	0.001
Pygoscelis adeliae	87	3872	81	3655	0.002	0.002
Pygoscelis antarctica	44	12,921	43	12320	0.002	0.003
Pygoscelis papua	36	4170	36	3496	0.013	0.001

Supplementary Table 16. *Aptenodytes* and *Pygoscelis* samples used for obtaining genome-wide SNPs. Genomic DNA is from Clucas *et al*. (2016), Younger *et al*. (2017) and Clucas *et al*. (2018). Note, we did not use the *P. papua* samples derived from Kerguelen, Falkland Islands or South Georgia in any analyses other than SNAPP because they exhibit population structure (see also *Figure 13*).

Taxon	Site name	Island group	Island Code	Collection seasons	Sample type	*n*
Aptenodytes patagonicus	Volunteer Pt.	Falkland Islands	1, FAL	2013 – 2014	Blood	16
Aptenodytes patagonicus	Fortuna Bay	South Georgia	5, SGE	2011 – 2012	Blood	16
Aptenodytes patagonicus	Baie du Marin	Crozet	10, CRZ	2003 – 2004	Blood	16
Aptenodytes patagonicus	Sandy Bay	Macquarie Island	13, MAC	2005 – 2006	Blood	16
Aptenodytes forsteri	Halley Bay	Weddell Sea, Antarctica	25, HAL	2012 – 2013	Tissue	13
Aptenodytes forsteri	Gould Bay	Weddell Sea, Antarctica	24, GOB	2013 – 2014	Blood	13
Aptenodytes forsteri	Fold Island	Antarctica	26, FOL	2010 – 2011	Tissue	16
Aptenodytes forsteri	Auster	Antarctica	28, AUS	1993 – 1994	Tissue	16
Aptenodytes forsteri	Amanda Bay	Antarctica	30, AMA	2012 – 2013	Tissue	16
Aptenodytes forsteri	Pointe Géologie	Antarctica	32, POI	2010 – 2011	Tissue	15
Aptenodytes forsteri	Cape Roget	Ross Sea, Antarctica	34, ROG	1992 – 1993	Blood	10
Aptenodytes forsteri	Cape Washington	Ross Sea, Antarctica	35, WAS	1992 – 1993	Blood	11
Pygoscelis antarctica	Nyrøysa	Bouvet Island	7, BOU	1997 – 1998	Blood	9
Pygoscelis antarctica	Hewison Point, Thule Island	South Sandwich Islands	6, SSI	2010 – 2011	Blood	10
Pygoscelis antarctica	Signy Island	South Orkney Islands	4, SOR	2006 – 2007	Blood	3
Pygoscelis antarctica	Admiralty Bay	South Shetland Islands	2, SSH	2010 – 2011	Blood	11
Pygoscelis antarctica	Orne Harbour	West Antarctic Peninsula	20, ORN	2013 – 2014	Blood	11
Pygoscelis adeliae	Bellingshausen Island	South Sandwich Islands	6, SSI	2010 – 2011	Blood	11
Pygoscelis adeliae	Signy Island	South Orkney Islands	4, SOR	2006 – 2007	Blood	3
Pygoscelis adeliae	Admiralty Bay	South Shetland Islands	2, SSH	2010 – 2011	Blood	11
Pygoscelis adeliae	Brown Bluff	North Antarctic Peninsula	23, BRO	1997 - 2000	Blood	11
Pygoscelis adeliae	Petermann Is.	West Antarctic Peninsula	19, PET	2013 – 2014	Blood	11
Pygoscelis adeliae	Béchervaise Island	Antarctica	27, BÉC	2012 – 2014	Tissue	11
Pygoscelis adeliae	Welch Island	Antarctica	29, WEL	2012 – 2014	Tissue	11
Pygoscelis adeliae	Blakeney Point	Antarctica	31, BLA	2012 – 2014	Tissue	10
Pygoscelis adeliae	Pétrels Island	Antarctica	33, PÉT	2012 – 2014	Tissue	9
Pygoscelis papua	Admiralty Bay	South Shetland Islands	2, SSH	2010 – 2011	Blood	11
Pygoscelis papua	George's Point	West Antarctic Peninsula	22, GEO	2013 – 2014	Blood	12
Pygoscelis papua	Jougla Point	West Antarctic Peninsula	21, JOU	2013 – 2014	Blood	11
Pygoscelis papua	Pointe du Morne	Kerguelen	11, KER	2013 – 2014	Blood	12
Pygoscelis papua	Cow Bay	Falkland Islands	1, FAL	2013 – 2014	Blood	11
Pygoscelis papua	South Georgia	Bird Island	5, SGE	2005 – 2006	Blood	10

Supplementary Table 17. Summary statistics calculated within each *Eudyptes* colony. The number of private alleles per each colony was also calculated across all *Eudyptes* species together (global analysis). In all cases, *Eudyptes chrysolophus chrysolophus* and *E. c. schlegeli* were analysed together. H_E = expected heterozygosity and H_O = observed heterozygosity. Note one individual each from Jackson Head and The Snares was removed in the global analysis during filtering.

Taxon	Site name	Island group	Island Code	*n* (global)	Number of private alleles (global)	Number of private alleles (species)	H_E	H_O	G_{IS}
Eudyptes pachyrhynchus	Codfish Island	New Zealand	18, COD	9	4	68	0.23	0.193	0.16
Eudyptes pachyrhynchus	Jackson Head	New Zealand	18, JAC	4 (3)	0	21	0.229	0.183	0.2
Eudyptes pachyrhynchus	Milford Sound	New Zealand	18, MIL	8	4	51	0.231	0.2	0.13
Eudyptes robustus	Snares	The Snares	17, SNA	4 (3)	6	233	0.378	0.259	0.32
Eudyptes robustus	Western Chain	Western Chain	17, WES	1	1	15	NA	NA	NA
Eudyptes chrysocome	New Island	Falkland Islands	1, NEW	13	1	68	0.184	0.152	0.17
Eudyptes chrysocome	Pebble Beach	Falkland Islands	1, PEB	13	5	81	0.204	0.17	0.17
Eudyptes moseleyi	Amsterdam Island	Amsterdam Island	12, AMS	6	11	216	0.303	0.237	0.22
Eudyptes moseleyi	Gough Island	Gough Island	8, GOU	2	1	36	0.323	0.245	0.24
Eudyptes filholi	Campbell Island	Campbell Island	14, CAM	13	8	40	0.119	0.098	0.18
Eudyptes filholi	Haskell Bay	Auckland Islands	15 AUC	8	5	6	0.118	0.098	0.17
Eudyptes filholi	Antipodes Islands	Antipodes Islands	16, ANT	3	1	5	0.116	0.099	0.15
Eudyptes filholi	Jardin Japonaise	Crozet	10, CRZ	15	0	11	0.123	0.108	0.15
Eudyptes filholi	Cave Bay	Prince Edward Island	9, PEI	20	6	41	0.123	0.108	0.12
Eudyptes filholi	Marion Island	Marion Island	9, MAR	4	0	3	0.126	0.103	0.18
Eudyptes filholi	Macquarie Island	Macquarie Island	13, MAC	8	0	11	0.122	0.1	0.18
Eudyptes chrysolophus schlegeli	Macquarie Island	Macquarie Island	13, MAC	35	42	54	0.07	0.059	0.16
Eudyptes chrysolophus ssp. 'white faced'	Marion Island	Marion Island	9, MAR	4	0	2	0.07	0.059	0.16
Eudyptes chrysolophus chrysolophus	Jardin Japonaise	Crozet	10, CRZ	17	16	23	0.067	0.056	0.16
Eudyptes chrysolophus chrysolophus	Prince Edward Island	Prince Edward Island	9, PEI	19	14	19	0.066	0.055	0.16
Eudyptes chrysolophus chrysolophus	South Georgia	South Georgia	5, SGE	13	12	18	0.064	0.053	0.16
Eudyptes chrysolophus chrysolophus	South Sandwich Islands	South Sandwich Islands	6, SSI	4	0	0	0.068	0.052	0.23
Eudyptes chrysolophus chrysolophus	Marion Island	Marion Island	9, MAR	11	3	4	0.069	0.055	0.2
Eudyptes chrysolophus chrysolophus	Kerguelen	Kerguelen	11, KER	7	3	7	0.071	0.06	157
Eudyptes chrysolophus chrysolophus	Seal Island	Elephant Island	3, ELE	15	7	15	0.068	0.058	0.14

Supplementary Table 18. Structure output for all *Eudyptes* datasets, as inferred by the Evanno method. All analyses were replicated twenty times.

Data	Location Priors	# *K*	Mean LnP (*K*)	Standard Deviation (*K*)	Ln' (*K*)	Ln'' (*K*)	Delta *K*
ALL *Eudyptes*	No	1	-973305.03	10.23	NA	NA	NA
ALL *Eudyptes*	No	2	-720549.75	6261.16	252755.29	131627.41	21.02
ALL *Eudyptes*	No	3	-599421.87	19.07	121127.88	101612.94	5328.91
ALL *Eudyptes*	No	4	-579906.94	42.08	19514.94	23208.31	551.59
ALL *Eudyptes*	No	5	-583600.31	15215.30	-3693.38	7681950.11	504.88
ALL *Eudyptes*	No	6	-8269243.80	13901981.78	-7685643.49	1113945.61	0.08
ALL *Eudyptes*	No	7	-17068832.89	21970543.03	-8799589.10	20153063.14	0.92
ALL *Eudyptes*	No	8	-5715358.85	14816861.06	11353474.05	9324249.00	0.63
ALL *Eudyptes*	No	9	-3686133.80	6720751.22	2029225.05	11846144.48	1.76
ALL *Eudyptes*	No	10	-13503053.23	22940908.21	-9816919.43	NA	NA
E. filholi, E. chrysocome, E. moseleyi	No	1	-406975.54	12.34	NA	NA	NA
E. filholi, E. chrysocome, E. moseleyi	No	2	-374907.87	1167.20	32067.67	2581.45	2.21
E. filholi, E. chrysocome, E. moseleyi	No	3	-345421.65	81.51	29486.22	31509.53	386.57
E. filholi, E. chrysocome, E. moseleyi	No	4	-347444.96	811.40	-2023.31	273082.49	336.56
E. filholi, E. chrysocome, E. moseleyi	No	5	-622550.75	426968.05	-275105.79	384739.23	0.90
E. filholi, E. chrysocome, E. moseleyi	No	6	-512917.31	320154.81	109633.44	321734.80	1.00
E. filholi, E. chrysocome, E. moseleyi	No	7	-725018.68	532345.80	-212101.37	44599.87	0.08
E. filholi, E. chrysocome, E. moseleyi	No	8	-981719.91	741555.06	-256701.23	147571.50	0.20
E. filholi, E. chrysocome, E. moseleyi	No	9	-1090849.64	924382.36	-109129.74	381199.18	0.41
E. filholi, E. chrysocome, E. moseleyi	No	10	-1581178.56	1668619.08	-490328.92	122061.20	0.07
E. filholi, E. chrysocome, E. moseleyi	No	11	-1949446.27	1013755.06	-368267.72	824678.60	0.81
E. filholi, E. chrysocome, E. moseleyi	No	12	-1493035.39	1536123.16	456410.88	405526.84	0.26
E. filholi, E. chrysocome, E. moseleyi	No	13	-1442151.35	1082049.88	50884.04	NA	NA
E. filholi, E. chrysocome, E. moseleyi	Yes	1	-406971.05	7.90	NA	NA	NA
E. filholi, E. chrysocome, E. moseleyi	Yes	2	-374805.49	1087.88	32165.56	2783.72	2.56
E. filholi, E. chrysocome, E. moseleyi	Yes	3	-345423.65	86.69	29381.84	32405.49	373.81
E. filholi, E. chrysocome, E. moseleyi	Yes	4	-348447.30	465.82	-3023.65	3320.81	7.13
E. filholi, E. chrysocome, E. moseleyi	Yes	5	-348150.14	1660.73	297.16	11791.01	7.10
E. filholi, E. chrysocome, E. moseleyi	Yes	6	-359643.99	40058.95	-11493.85	191578.51	4.78
E. filholi, E. chrysocome, E. moseleyi	Yes	7	-562716.34	390414.96	-203072.35	29394.02	0.08
E. filholi, E. chrysocome, E. moseleyi	Yes	8	-736394.67	405679.25	-173678.34	65734.68	0.16
E. filholi, E. chrysocome, E. moseleyi	Yes	9	-844338.33	515617.23	-107943.66	283187.83	0.55
E. filholi, E. chrysocome, E. moseleyi	Yes	10	-1235469.82	910202.35	-391131.49	170465.43	0.19
E. filholi, E. chrysocome, E. moseleyi	Yes	11	-1797066.74	1087928.47	-561596.92	956124.33	0.88
E. filholi, E. chrysocome, E. moseleyi	Yes	12	-1402539.33	800249.47	394527.41	613102.48	0.77
E. filholi, E. chrysocome, E. moseleyi	Yes	13	-1621114.40	1041024.72	-218575.07	NA	NA

Data	Location Priors	# *K*	Mean LnP (*K*)	Standard Deviation (*K*)	Ln' (*K*)	Ln'' (*K*)	Delta *K*
E. filholi, E. chrysocome	No	1	-392521.21	7.66	NA	NA	NA
E. filholi, E. chrysocome	No	2	-360089.70	95.86	32431.52	33487.63	349.34
E. filholi, E. chrysocome	No	3	-361145.81	87.96	-1056.11	70413.30	800.54
E. filholi, E. chrysocome	No	4	-432615.21	193277.34	-71469.41	97837.18	0.51
E. filholi, E. chrysocome	No	5	-406247.44	135077.46	26367.78	414799.70	3.07
E. filholi, E. chrysocome	No	6	-794679.36	582934.42	-388431.93	250552.82	0.43
E. filholi, E. chrysocome	No	7	-932558.47	619842.95	-137879.11	88176.74	0.14
E. filholi, E. chrysocome	No	8	-1158614.32	711913.40	-226055.85	42882.41	0.06
E. filholi, E. chrysocome	No	9	-1427552.58	1280606.20	-268938.26	103571.27	0.08
E. filholi, E. chrysocome	No	10	-1592919.56	1118962.27	-165366.99	398831.35	0.36
E. filholi, E. chrysocome	No	11	-2157117.90	1476919.28	-564198.34	NA	NA
E. filholi, E. chrysocome	Yes	1	-392526.06	8.16	NA	NA	NA
E. filholi, E. chrysocome	Yes	2	-359927.37	46.03	32598.69	33588.51	729.79
E. filholi, E. chrysocome	Yes	3	-360917.18	64.93	-989.82	1667.46	25.68
E. filholi, E. chrysocome	Yes	4	-363574.45	4966.44	-2657.27	4092.17	0.82
E. filholi, E. chrysocome	Yes	5	-362139.56	11579.76	1434.90	363199.56	31.37
E. filholi, E. chrysocome	Yes	6	-723904.22	473485.70	-361764.67	56312.14	0.12
E. filholi, E. chrysocome	Yes	7	-1029356.75	609208.88	-305452.53	47372.39	0.08
E. filholi, E. chrysocome	Yes	8	-1287436.90	833067.54	-258080.15	22528.58	0.03
E. filholi, E. chrysocome	Yes	9	-1522988.47	707175.83	-235551.57	305735.44	0.43
E. filholi, E. chrysocome	Yes	10	-2064275.48	897267.02	-541287.01	31849.67	0.04
E. filholi, E. chrysocome	Yes	11	-2573712.82	1322006.68	-509437.34	NA	NA
E. filholi	No	1	-270683.09	6.40	NA	NA	NA
E. filholi	No	2	-270066.42	21.55	616.67	10977.54	509.45
E. filholi	No	3	-280427.29	40635.75	-10360.87	6671.83	0.16
E. filholi	No	4	-297459.99	62549.43	-17032.70	11034.10	0.18
E. filholi	No	5	-325526.79	152014.11	-28066.80	76934.91	0.51
E. filholi	No	6	-430528.49	279318.48	-105001.71	67992.34	0.24
E. filholi	No	7	-467537.86	481242.92	-37009.37	33241.20	0.07
E. filholi	No	8	-471306.04	398710.14	-3768.18	94194.58	0.24
E. filholi	No	9	-380879.64	194143.12	90426.40	NA	NA
E. filholi	Yes	1	-270682.97	7.38	NA	NA	NA
E. filholi	Yes	2	-270078.13	27.40	604.84	975.63	35.60
E. filholi	Yes	3	-270448.92	211.05	-370.79	20662.60	97.91
E. filholi	Yes	4	-291482.31	45220.54	-21033.39	43683.65	0.97
E. filholi	Yes	5	-268832.05	1739.51	22650.26	417267.73	239.88
E. filholi	Yes	6	-663449.53	567904.67	-394617.48	572195.86	1.01
E. filholi	Yes	7	-485871.14	229554.02	177578.39	42740.85	0.19

Data	Location Priors	# *K*	Mean LnP (*K*)	Standard Deviation (*K*)	Ln' (*K*)	Ln'' (*K*)	Delta *K*
E. filholi	Yes	8	-351033.61	161033.19	134837.54	458658.20	2.85
E. filholi	Yes	9	-674854.27	481447.23	-323820.66	NA	NA
E. chrysocome	No	1	-128191.88	21.21	NA	NA	NA
E. chrysocome	No	2	-126760.41	41.71	1431.47	13319.55	319.36
E. chrysocome	No	3	-138648.48	36717.51	-11888.08	20420.76	0.56
E. chrysocome	No	4	-130115.80	18385.48	8532.69	NA	NA
E. chrysocome	Yes	1	-128210.26	252.87	NA	NA	NA
E. chrysocome	Yes	2	-126768.09	34.67	1442.17	9538.79	275.15
E. chrysocome	Yes	3	-134864.71	26748.93	-8096.62	3487.66	0.13
E. chrysocome	Yes	4	-146448.99	32228.03	-11584.28	NA	NA
E. moseleyi	No	1	-36950.24	24.88	NA	NA	NA
E. moseleyi	No	2	-37557.32	468.20	-607.09	2341.00	5.00
E. moseleyi	No	3	-40505.40	2349.77	-2948.08	527.62	0.22
E. moseleyi	No	4	-42925.86	5199.38	-2420.46	NA	NA
E. moseleyi	Yes	1	-36951.71	22.07	NA	NA	NA
E. moseleyi	Yes	2	-37627.02	336.29	-675.32	809.80	2.41
E. moseleyi	Yes	3	-39112.13	783.02	-1485.11	837.43	1.07
E. moseleyi	Yes	4	-41434.67	5196.25	-2322.54	NA	NA
E. c. chrysolophus, E. c. schlegeli	No	1	-429760.63	10.34	NA	NA	NA
E. c. chrysolophus, E. c. schlegeli	No	2	-428300.46	34.70	1460.17	2433.60	70.12
E. c. chrysolophus, E. c. schlegeli	No	3	-429273.90	216.09	-973.44	6411.03	29.67
E. c. chrysolophus, E. c. schlegeli	No	4	-436658.36	14119.99	-7384.46	66.44	0.00
E. c. chrysolophus, E. c. schlegeli	No	5	-443976.38	35845.81	-7318.03	188191.99	5.25
E. c. chrysolophus, E. c. schlegeli	No	6	-639486.39	325625.50	-195510.01	8064.03	0.02
E. c. chrysolophus, E. c. schlegeli	No	7	-843060.43	485929.56	-203574.04	335028.77	0.69
E. c. chrysolophus, E. c. schlegeli	No	8	-711605.69	537618.95	131454.74	403603.55	0.75
E. c. chrysolophus, E. c. schlegeli	No	9	-983754.50	700505.69	-272148.81	377927.31	0.54
E. c. chrysolophus, E. c. schlegeli	No	10	-1633830.62	888784.16	-650076.12	825865.82	0.93
E. c. chrysolophus, E. c. schlegeli	No	11	-1458040.92	788179.37	175789.70	NA	NA
E. c. chrysolophus, E. c. schlegeli	Yes	1	-429763.08	13.57	NA	NA	NA
E. c. chrysolophus, E. c. schlegeli	Yes	2	-428303.38	30.66	1459.70	3170.28	103.41
E. c. chrysolophus, E. c. schlegeli	Yes	3	-430013.96	824.63	-1710.58	7555.29	9.16
E. c. chrysolophus, E. c. schlegeli	Yes	4	-439279.83	19225.86	-9265.87	74640.97	3.88
E. c. chrysolophus, E. c. schlegeli	Yes	5	-523186.66	128655.92	-83906.84	47680.38	0.37
E. c. chrysolophus, E. c. schlegeli	Yes	6	-559413.12	150997.46	-36226.46	158962.26	1.05
E. c. chrysolophus, E. c. schlegeli	Yes	7	-754601.83	518787.90	-195188.71	1755.37	0.00
E. c. chrysolophus, E. c. schlegeli	Yes	8	-948035.17	586353.82	-193433.35	30085.08	0.05
E. c. chrysolophus, E. c. schlegeli	Yes	9	-1171553.59	940540.60	-223518.42	15495.54	0.02

Data	Location Priors	# *K*	Mean LnP (*K*)	Standard Deviation (*K*)	Ln' (*K*)	Ln'' (*K*)	Delta *K*
E. c. chrysolophus, E. c. schlegeli	Yes	10	-1379576.48	813042.71	-208022.89	71320.07	0.09
E. c. chrysolophus, E. c. schlegeli	Yes	11	-1516279.30	600051.36	-136702.82	NA	NA
E. pachyrhynchus, E. robustus	No	1	-79665.25	59.88	NA	NA	NA
E. pachyrhynchus, E. robustus	No	2	-69048.13	8.36	10617.12	12065.63	1442.61
E. pachyrhynchus, E. robustus	No	3	-70496.65	76.86	-1448.52	453.60	5.90
E. pachyrhynchus, E. robustus	No	4	-72398.76	5070.63	-1902.12	1699.37	0.34
E. pachyrhynchus, E. robustus	No	5	-72601.51	4444.60	-202.75	2968.56	0.67
E. pachyrhynchus, E. robustus	No	6	-75772.81	6133.13	-3171.30	14933.23	2.43
E. pachyrhynchus, E. robustus	No	7	-93877.34	28767.44	-18104.53	NA	NA
E. pachyrhynchus, E. robustus	Yes	1	-79638.41	77.30	NA	NA	NA
E. pachyrhynchus, E. robustus	Yes	2	-69045.67	8.61	10592.74	11910.82	1383.35
E. pachyrhynchus, E. robustus	Yes	3	-70363.75	171.67	-1318.08	2183.20	12.72
E. pachyrhynchus, E. robustus	Yes	4	-73865.02	5609.28	-3501.28	2330.95	0.42
E. pachyrhynchus, E. robustus	Yes	5	-75035.35	6406.13	-1170.33	1222.34	0.19
E. pachyrhynchus, E. robustus	Yes	6	-74983.34	4010.40	52.01	19297.06	4.81
E. pachyrhynchus, E. robustus	Yes	7	-94228.39	27187.25	-19245.05	NA	NA
E. pachyrhynchus	No	1	-67225.11	14.81	NA	NA	NA
E. pachyrhynchus	No	2	-77282.25	9782.69	-10057.14	12703.11	1.30
E. pachyrhynchus	No	3	-74636.27	14364.02	2645.98	377.86	0.03
E. pachyrhynchus	No	4	-72368.16	7968.42	2268.12	3592.51	0.45
E. pachyrhynchus	No	5	-73692.55	4283.91	-1324.39	NA	NA
E. pachyrhynchus	Yes	1	-67222.95	14.97	NA	NA	NA
E. pachyrhynchus	Yes	2	-71809.76	7135.83	-4586.82	4706.71	0.66
E. pachyrhynchus	Yes	3	-71689.87	15327.87	119.89	6883.42	0.45
E. pachyrhynchus	Yes	4	-78453.40	13671.63	-6763.53	24656.62	1.80
E. pachyrhynchus	Yes	5	-109873.55	30237.09	-31420.15	NA	NA
E. robustus	No	1	-17830.75	16.10	NA	NA	NA
E. robustus	No	2	-17317.50	65.12	513.26	2697.73	41.43
E. robustus	No	3	-19501.97	4918.13	-2184.48	885.07	0.18
E. robustus	No	4	-20801.38	4887.89	-1299.41	25796.59	5.28
E. robustus	No	5	-47897.38	75351.85	-27096.00	NA	NA
E. robustus	Yes	1	-17825.63	17.53	NA	NA	NA
E. robustus	Yes	2	-18569.95	3724.40	-744.32	206.94	0.06
E. robustus	Yes	3	-19107.33	5550.66	-537.38	1962.27	0.35
E. robustus	Yes	4	-21606.98	7878.78	-2499.65	26945.29	3.42
E. robustus	Yes	5	-51051.91	47733.27	-29444.94	NA	NA

Supplementary Table 19. Samples used in thirteen independent SNAPP analyses. *Eudyptes* penguins were run in species clusters; (1) *E. chrysocome*, *E. filholi* and *E. moseleyi*; (2) *E. chrysolophus chrysolophus* and *E. c. schlegeli*; (3) *E. pachyrhynchus* and *E. robustus* for three replicates each (see *Supplementary Figure 15*), and the *Pygoscelis* penguins were run as independent species; *P. adeliae* and *P. antarctica*, for two replicates each (see *Supplementary Figure 14*). Each replicate included two different individuals from each site. SNAPP analyses for *P. papua*, *Aptenodytes patagonicus* and *A. forsteri* are published in Clucas *et al*. (2016), Younger *et al*. (2017) and Clucas *et al*. (2018). Site details are follows: PEB is Pebble Beach (Falkland Islands); NEW is New Island (Falkland Islands); MAC is Macquarie Island; CAM is Campbell Island; AUC is Auckland Islands; ANT is Antipodes Islands; CRZ is Crozet; PEI is Prince Edward Islands; MAR is Marion Island; AMS is Amsterdam Island; GOU is Gough Island; KER is Kerguelen; ELE is Elephant Island; SGE is South Georgia; SSI is South Sandwich Islands; MAR* is 'white faced' penguins from Marion Island; COD is Codfish Island; MIL is Milford Sound; JAC is Jackson Head; SNA is The Snares; WES is Western Chain; PET is Peterman Island; BRO is Brown Bluff; SSH is South Shetland Islands; SOR is South Orkney Islands; BÉC is Béchervaise Island; WEL is Welch Island; BLA is Blakeney Point; PÉT is Pétrels Island; ORN is Orne Harbour; BOU is Bouvet.

Taxon	Site	Rep 1	Rep 2	Rep 3
Eudyptes chrysocome, Eudyptes filholi, Eudyptes moseleyi				
E. chrysocome	**PEB**	JW1082, JW1083	JW1084, JW1085	JW1086, JW1087
E. chrysocome	**NEW**	JW722, JW724	JW726, JW728	JW730, JW731
E. filholi	**MAC**	JW1195, JW1196	JW1197, JW1197	JW1199, JW1200
E. filholi	**CAM**	JW531, JW532	JW569, JW537	JW542, JW543
E. filholi	**AUC**	JW677, JW678	JW679/700, JW680/701	JW684, JW685
E. filholi	**ANT**	JW691, JW692	JW690, JW692	JW690, JW691
E. filholi	**CRZ**	JW744, JW745	JW746, JW747	JW748, JW749
E. filholi	**PEI**	JW798, JW799	JW800, JW840	JW801, JW802
E. filholi	**MAR**	JW968, JW975	JW960, JW968	JW965, JW968
E. moseleyi	**AMS**	JW1046/1102, JW1048	JW1049, JW1057	JW1044, JW1049
E. moseleyi	**GOU**	JW1123, JW1124	JW1123, JW1124	JW1123, JW1124
Eudyptes chrysolophus chrysolophus/Eudyptes chrysolophus schlegeli				
E. c. chrysolophus	**KER**	JW1101, JW1016/1073	JW1019, JW1020	JW1006, JW1022
E. c. chrysolophus	**ELE**	JW1028, JW1030	JW1031, JW1032	JW1033, JW1034
E. c. chrysolophus	**CRZ**	JW761, JW764	JW765, JW766	JW767, JW768
E. c. chrysolophus	**PEI**	JW823, JW865	JW825, JW826	JW827, JW828/866
E. c. chrysolophus	**SGE**	JW847, JW849	JW850, JW852	JW853, JW855
E. c. chrysolophus	**SSI**	JW932, JW942	JW928, JW932	JW931, JW942
E. c. chrysolophus	**MAR**	JW947, JW948	JW928, JW932	JW951, JW953
***E. chrysolophus* hybrids**	**MAR***	JW792, JW794	JW782, JW792	JW785, JW794
E. c. schlegeli	**MAC**	JW871, JW872	JW889, JW890	JW874, JW892
Eudyptes pachyrhynchus/ Eudyptes robustus				
E. pachyrhynchus	**COD**	JW316/643, JW320/647	JW638, JW639	JW645, JW646/912
E. pachyrhynchus	**MIL**	JW439, JW661	JW663, JW664	JW343, JW924
E. pachyrhynchus	**JAC**	JW655, JW660/921	JW654, JW656/918	JW655, JW659
E. robustus	**SNA**	JW480/708, JW713	JW480/708, JW482/710	JW712, JW713
E. robustus	**WES**	JW485	JW485	JW485
Pygoscelis adeliae				
P. adeliae	**PET**	APPI075, APPI078	APPI071, APPI072	NA
P. adeliae	**BRO**	APBR007, APBR017	APBR003, APBR020	NA
P. adeliae	**SSH**	APKG022, APKG025	APKG021, APKG028	NA
P. adeliae	**SOR**	APSG031, APSG032	APSG032, APSG033	NA
P. adeliae	**SSI**	APBS007, APBS008	APBS006, APBS019	NA
P. adeliae	**BÉC**	APBI014, APBI017	APBI007, APBI024	NA
P. adeliae	**WEL**	APWH003, APWH022	APWH018, APWH030	NA
P. adeliae	**BLA**	APBP011, APBP018	APBP015, APBP016	NA
P. adeliae	**PÉT**	APAL021, APAL026	APAL024, APAL032	NA
Pygoscelis Antarctica				
P. antarctica	**ORN**	CPOH106, CPOH112	CPOH103. CPOH109	NA
P. antarctica	**SSH**	CPKG069, CPKG070	CPKG075, CPKG080	NA
P. antarctica	**SOR**	CPSG051, CPSG052	CPSG051, CPSG053	NA
P. antarctica	**SSI**	CPTH052, CPTH055	CPTH051, CPTH064	NA
P. antarctica	**BOU**	CPBV004, CPBV007	CPBV008, CPBV009	NA

Supplementary Table 20. F_{ST} between all sampled colonies of *Eudyptes* penguins. F_{ST} is below, corrected *P* value are above. *P* values <0.05 are shown in bold (all values except nine). All analyses were calculated across 10,000 bootstraps. Site codes are as follows: NEW is New Island (Falkland Islands); PEB is Pebble Beach (Falkland Islands); CAM is Campbell Island; AUC is Auckland Islands; ANT is Antipodes Islands; CRZ is Crozet; PEI is Prince Edward Islands; MAR is Marion Island; MAC is Macquarie Island; AMS is Amsterdam Island; GOU is Gough Island; ELE is Elephant Island; SGE is South Georgia; SSI is South Sandwich Islands; KER is Kerguelen; MAR* is 'white-faced' penguins from Marion Island; COD is Codfish Island; MIL is Milford Sound; JAC is Jackson Head; SNA is The Snares and WES is Western Chain.

		E. chrysocome NEW	*E. chrysocome* PEB	*E. filholi* CAM	*E. filholi* AUC	*E. filholi* ANT	*E. filholi* CRZ	*E. filholi* PEI	*E. filholi* MAR	*E. filholi* MAC	*E. moseleyi* AMS	*E. moseleyi* GOU	*E. c. chrysolophus* ELE	*E. c. chrysolophus* SGE	*E. c. chrysolophus* SSI	*E. c. chrysolophus* PEI	*E. c. chrysolophus* MAR	*E. c. chrysolophus* CRZ	*E. c. chrysolophus* KER	*E. chrysolophus* spp. MAR*	*E. c. schlegeli* MAC	*E. pachyrhynchus* COD	*E. pachyrhynchus* MIL	*E. pachyrhynchus* JAC	*E. robustus* SNA	*E. robustus* WES
E. chrysocome	**NEW**	-	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**
E. chrysocome	**PEB**	0.008	-	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**
E. filholi	**CAM**	0.174	0.157	-	**0.000**	0.777	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**
E. filholi	**AUC**	0.187	0.165	0.015	-	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**
E. filholi	**ANT**	0.210	0.183	0.001	0.015	-	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**
E. filholi	**CRZ**	0.172	0.155	0.030	0.027	0.030	-	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**
E. filholi	**PEI**	0.173	0.159	0.032	0.032	0.035	0.011	-	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**
E. filholi	**MAR**	0.202	0.179	0.031	0.028	0.032	0.007	0.009	-	**0.002**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**
E. filholi	**MAC**	0.192	0.169	0.028	0.027	0.024	0.011	0.013	0.011	-	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**
E. moseleyi	**AMS**	0.457	0.433	0.038	0.389	0.430	0.354	0.351	0.373	0.388	-	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**
E. moseleyi	**GOU**	0.433	0.398	0.332	0.333	0.361	0.304	0.309	0.312	0.312	0.045	-	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**
E. c. chrysolophus	**ELE**	0.548	0.534	0.512	0.517	0.534	0.504	0.501	0.522	0.515	0.582	0.565	-	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**
E. c. chrysolophus	**SGE**	0.556	0.540	0.516	0.523	0.543	0.507	0.503	0.530	0.520	0.593	0.577	0.011	-	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**
E. c. chrysolophus	**SSI**	0.560	0.536	0.495	0.498	0.521	0.486	0.485	0.497	0.492	0.600	0.568	0.015	0.011	-	**0.000**	**0.022**	**0.000**	**0.024**	**0.000**	**0.005**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**
E. c. chrysolophus	**PEI**	0.551	0.538	0.519	0.524	0.540	0.511	0.508	0.529	0.522	0.585	0.569	0.019	0.015	0.485	-	0.440	**0.000**	**0.022**	**0.008**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**
E. c. chrysolophus	**MAR**	0.552	0.536	0.510	0.514	0.532	0.501	0.499	0.518	0.511	0.585	0.564	0.019	0.011	0.012	0.002	-	**0.000**	0.441	0.510	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**
E. c. chrysolophus	**CRZ**	0.553	0.539	0.518	0.523	0.540	0.510	0.528	0.507	0.521	0.587	0.570	0.021	0.014	0.018	0.005	0.006	-	**0.002**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**
E. c. chrysolophus	**KER**	0.553	0.534	0.502	0.506	0.524	0.493	0.491	0.508	0.502	0.589	0.563	0.020	0.014	0.012	0.006	0.010	0.007	-	**0.002**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**

		E. robustus WES	*E. robustus* SNA	*E. pachyrhynchus* JAC	*E. pachyrhynchus* MIL	*E. pachyrhynchus* COD	*E. c. schlegeli* MAC	*E. chrysolophus* spp. MAR*	*E. c. chrysolophus* KER	*E. c. chrysolophus* CRZ	*E. c. chrysolophus* MAR	*E. c. chrysolophus* PEI	*E. c. chrysolophus* SSI	*E. c. chrysolophus* SGE	*E. c. chrysolophus* ELE	*E. moseleyi* GOU	*E. moseleyi* AMS	*E. filholi* MAC	*E. filholi* MAR	*E. filholi* PEI	*E. filholi* CRZ	*E. filholi* ANT	*E. filholi* AUC	*E. filholi* CAM	*E. chrysocome* PEB	*E. chrysocome* NEW
E. chrysolophus ssp.	MAR*	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	-	0.011	0.011	0.005	0.011	0.022	0.021	0.022	0.569	0.597	0.491	0.525	0.511	0.485	0.519	0.498	0.495	0.535	0.559
E. c. schlegeli	MAC	**0.000**	**0.000**	**0.000**	**0.000**	-	-	- 0.003	0.013	0.013	0.010	0.011	0.022	0.017	0.023	0.559	0.573	0.522	0.525	0.511	0.513	0.534	0.522	0.519	0.531	0.540
E. pachyrhynchus	COD	**0.000**	**0.000**	0.459	0.441	-	0.662	0.693	0.683	0.676	0.678	0.675	0.694	0.683	0.674	0.689	0.695	0.626	0.647	0.603	0.608	0.663	0.634	0.620	0.654	0.671
E. pachyrhynchus	MIL	**0.000**	**0.000**	0.639	-	0.005	0.661	0.690	0.680	0.674	0.676	0.673	0.692	0.681	0.672	0.689	0.695	0.623	0.643	0.600	0.605	0.661	0.631	0.617	0.653	0.670
E. pachyrhynchus	JAC	**0.000**	**0.000**	-	0.009	0.005	0.662	0.675	0.664	0.664	0.653	0.663	0.677	0.671	0.661	0.688	0.693	0.593	0.608	0.581	0.583	0.635	0.604	0.595	0.641	0.583
E. robustus	SNA	-	0.312	0.316	0.312	0.331	0.617	0.628	0.619	0.628	0.621	0.637	0.624	0.635	0.623	0.652	0.657	0.522	0.517	0.523	0.521	0.548	0.534	0.534	0.597	0.627
E. robustus	WES	0.535	-	0.319	0.335	0.331	0.617	0.628	0.619	0.628	0.621	0.637	0.624	0.635	0.623	0.639	0.657	0.522	0.517	0.523	0.521	0.548	0.534	0.534	0.597	0.627

Supplementary Table 21. F_{ST} between all sampled colonies of *Eudyptes chrysolophus chrysolophus* and *E. c. schlegeli*. F_{ST} is below, corrected *P* value are above. *P* values <0.05 are shown in bold (all values except three). All analyses were calculated across 10,000 bootstraps. Site codes are as follows: ELE is Elephant Island; SGE is South Georgia; SSI is South Sandwich Islands; PEI is Prince Edward Islands; MAR is Marion Island; MAR* are 'white-faced' penguins from Marion Island; CRZ is Crozet; KER is Kerguelen; and MAC is Macquarie Island.

		Eudyptes chrysolophus chrysolophus ELE	*Eudyptes chrysolophus chrysolophus* SGE	*Eudyptes chrysolophus chrysolophus* SSI	*Eudyptes chrysolophus chrysolophus* PEI	*Eudyptes chrysolophus chrysolophus* MAR	*Eudyptes chrysolophus* ssp. MAR*	*Eudyptes chrysolophus chrysolophus* CRZ	*Eudyptes chrysolophus chrysolophus* KER	*Eudyptes chrysolophus schlegeli* MAC
Eudyptes chrysolophus chrysolophus	**ELE**	-	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**
Eudyptes chrysolophus chrysolophus	**SGE**	0.010	-	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**
Eudyptes chrysolophus chrysolophus	**SSI**	0.015	0.014	-	**0.000**	**0.020**	**0.006**	**0.000**	**0.011**	**0.000**
Eudyptes chrysolophus chrysolophus	**PEI**	0.018	0.011	0.020	-	**0.076**	**0.000**	**0.000**	**0.000**	**0.000**
Eudyptes chrysolophus chrysolophus	**MAR**	0.017	0.012	0.009	0.002	-	0.075	**0.000**	0.595	**0.000**
***Eudyptes chrysolophus* ssp.**	**MAR***	0.022	0.021	0.015	0.012	0.006	-	**0.000**	**0.000**	0.917
Eudyptes chrysolophus chrysolophus	**CRZ**	0.021	0.016	0.022	0.005	0.006	0.014	-	**0.000**	**0.000**
Eudyptes chrysolophus chrysolophus	**KER**	0.016	0.014	0.011	0.006	0.001	0.009	0.007	-	**0.016**
Eudyptes chrysolophus schlegeli	**MAC**	0.022	0.017	0.021	0.011	0.010	-0.003	0.013	0.012	-

Supplementary Table 22. F_{ST} between all sampled colonies of *Eudyptes chrysocome*, *E. filholi* and *E. moseleyi*. F_{ST} is below, corrected *P* value are above. *P* values <0.05 are shown in bold (all values except one). All analyses were calculated across 10,000 bootstraps. Site codes are as follows: NEW is New Island (Falkland Islands); PEB is Pebble Beach (Falkland Islands); CAM is Campbell Island; AUC is Auckland Islands; ANT is Antipodes Islands; CRZ is Crozet; PEI is Prince Edward Islands; MAR is Marion Island; MAC is Macquarie Island; AMS is Amsterdam Island; and GOU is Gough Island.

		Eudyptes chrysocome NEW	*Eudyptes chrysocome* PEB	*Eudyptes filholi* CAM	*Eudyptes filholi* AUC	*Eudyptes filholi* ANT	*Eudyptes filholi* CRZ	*Eudyptes filholi* PEI	*Eudyptes filholi* MAR	*Eudyptes filholi* MAC	*Eudyptes moseleyi* AMS	*Eudyptes moseleyi* GOU
Eudyptes chrysocome	**NEW**	-	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**
Eudyptes chrysocome	**PEB**	0.006	-	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**
Eudyptes filholi	**CAM**	0.175	0.161	-	**0.000**	0.829	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**
Eudyptes filholi	**AUC**	0.185	0.168	0.014	-	**0.011**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**
Eudyptes filholi	**ANT**	0.206	0.181	-0.003	0.010	-	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**
Eudyptes filholi	**CRZ**	0.177	0.164	0.034	0.031	0.269	-	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**
Eudyptes filholi	**PEI**	0.173	0.162	0.035	0.032	0.030	0.008	-	**0.000**	**0.000**	**0.000**	**0.000**
Eudyptes filholi	**MAR**	0.199	0.180	0.035	0.032	0.029	0.012	0.010	-	**0.000**	**0.000**	**0.000**
Eudyptes filholi	**MAC**	0.189	0.170	0.030	0.026	0.157	0.010	0.010	0.011	-	**0.000**	**0.000**
Eudyptes moseleyi	**AMS**	0.487	0.467	0.414	0.422	0.458	0.385	0.378	0.413	0.402	-	**0.000**
Eudyptes moseleyi	**GOU**	0.467	0.438	0.372	0.372	0.402	0.342	0.341	0.342	0.347	0.055	-

Supplementary Table 23. F_{ST} between all sampled colonies of *Eudyptes chrysocome* and *E. filholi*. F_{ST} is below, corrected *P* value are above. *P* values <0.05 are shown in bold (all values except one). All analyses were calculated across 10,000 bootstraps. Site codes are as follows: NEW is New Island (Falkland Islands); PEB is Pebble Beach (Falkland Islands); CAM is Campbell Island; AUC is Auckland Islands; ANT is Antipodes Islands; CRZ is Crozet; PEI is Prince Edward Islands; MAR is Marion Island; and MAC is Macquarie Island.

		Eudyptes chrysocome NEW	*Eudyptes chrysocome* PEB	*Eudyptes filholi* CAM	*Eudyptes filholi* AUC	*Eudyptes filholi* ANT	*Eudyptes filholi* CRZ	*Eudyptes filholi* PEI	*Eudyptes filholi* MAR	*Eudyptes filholi* MAC
Eudyptes chrysocome	**NEW**	-	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**
Eudyptes chrysocome	**PEB**	0.006	-	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**
Eudyptes filholi	**CAM**	0.178	0.165	-	**0.000**	0.639	**0.000**	**0.000**	**0.000**	**0.000**
Eudyptes filholi	**AUC**	0.189	0.172	0.015	-	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**
Eudyptes filholi	**ANT**	0.209	0.185	-0.001	0.012	-	**0.000**	**0.000**	**0.000**	**0.000**
Eudyptes filholi	**CRZ**	0.179	0.166	0.335	0.030	0.028	-	**0.000**	**0.000**	**0.000**
Eudyptes filholi	**PEI**	0.176	0.164	0.035	0.032	0.031	0.008	-	**0.000**	**0.000**
Eudyptes filholi	**MAR**	0.203	0.183	0.033	0.030	0.029	0.009	0.010	-	**0.007**
Eudyptes filholi	**MAC**	0.193	0.174	0.025	0.025	0.018	0.009	0.010	0.008	-

Supplementary Table 24. F_{ST} between all sampled colonies of *Eudyptes filholi*. F_{ST} is below, corrected *P* value are above. *P* values <0.05 are shown in bold (all values except one). All analyses were calculated across 10,000 bootstraps. Site codes are as follows: CAM is Campbell Island; AUC is Auckland Islands; ANT is Antipodes Islands; CRZ is Crozet; PEI is Prince Edward Islands; MAR is Marion Island; and MAC is Macquarie Island.

	CAM	**AUC**	**ANT**	**CRZ**	**PEI**	**MAR**	**MAC**
CAM	-	**0.000**	0.257	**0.000**	**0.000**	**0.000**	**0.000**
AUC	0.013	-	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**
ANT	0.002	0.013	-	**0.000**	**0.000**	**0.000**	**0.000**
CRZ	0.032	0.028	0.030	-	**0.000**	**0.000**	**0.000**
PEI	0.034	0.031	0.033	0.009	-	**0.000**	**0.000**
MAR	0.032	0.030	0.031	0.010	0.011	-	**0.005**
MAC	0.029	0.025	0.023	0.011	0.011	0.009	-

Supplementary Table 25. F_{ST} between all sampled colonies of *Eudyptes chrysocome*. F_{ST} is below, corrected *P* value is above. The *P* value is <0.05 (shown in bold). All analyses were calculated across 10,000 bootstraps. Site codes are as follows: NEW is New Island and PEB is Pebble Beach.

	NEW	**PEB**
NEW	-	**0.000**
PEB	0.006	-

Supplementary Table 26. F_{ST} between all sampled colonies of *Eudyptes moseleyi*. F_{ST} is below, corrected *P* value is above. The *P* value is <0.05 (shown in bold). All analyses were calculated across 10,000 bootstraps. Site codes are as follows: AMS is Amsterdam Island and GOU is Gough Island.

	AMS	**GOU**
AMS	-	**0.000**
GOU	0.062	-

Supplementary Table 27. F_{ST} between all sampled colonies of *Eudyptes pachyrhynchus* and *E. robustus*. F_{ST} is below, corrected *P* value are above. *P* values <0.05 are shown in bold (all values except one). All analyses were calculated across 10,000 bootstraps. Site codes are as follows: COD is Codfish Island MIL is Milford Sound; JAC is Jackson Head; SNA is The Snares; and WES is Western Chain.

		Eudyptes pachyrhynchus COD	*Eudyptes pachyrhynchus* MIL	*Eudyptes pachyrhynchus* JAC	*Eudyptes robustus* SNA	*Eudyptes robustus* WES
Eudyptes pachyrhynchus	**COD**	-	**0.020**	**0.020**	**0.000**	**0.000**
Eudyptes pachyrhynchus	**MIL**	0.005	-	**0.020**	**0.000**	**0.000**
Eudyptes pachyrhynchus	**JAC**	0.008	0.009	-	**0.000**	**0.000**
Eudyptes robustus	**SNA**	0.324	0.320	0.320	-	0.241
Eudyptes robustus	**WES**	0.337	0.336	0.328	0.010	-

Supplementary Table 28. F_{ST} between all sampled colonies of *Eudyptes pachyrhynchus*. F_{ST} is below, corrected *P* value is above. The *P* values are <0.05 (shown in bold). All analyses were calculated across 10,000 bootstraps. Site codes are as follows: COD is Codfish Island; MIL is Milford Sound; and JAC is Jackson Head.

	COD	**MIL**	**JAC**
COD	-	**0.014**	**0.002**
MIL	0.005	-	**0.014**
JAC	0.010	0.008	-

Supplementary Table 29. F_{ST} between all sampled colonies of *Eudyptes robustus*. F_{ST} is below, corrected *P* value is above. The *P* value is >0.05. All analyses were calculated across 10,000 bootstraps. Site codes are as follows: SNA is The Snares; WES is Western Chain.

	SNA	**WES**
SNA	-	0.280
WES	0.007	-

Supplementary Table 30. Summary of the demographic histories of *Eudyptes*, *Pygoscelis* and *Aptendoytes* as inferred by CubSFS. The summary includes whether CubSFS infers an expansion, the time of the expansion, the N_e at the start of expansion, the N_e at time 1000 years BCE, the current N_e, and the fold increase. Results are based on the observed values. Note *E. c. chrysolophus* and *E. c. schlegeli* samples were run in the same analysis. Note also that *E. robustus* samples were not analysed using CubSFS due to small sample sizes.

Taxon	Expansion	Time at start of expansion	N_e at start of expansion	N_e 1000 BCE	Current N_e	Fold increase
Aptenodytes forsteri	Yes	40000	59361.76	224129.63	261390.18	4.40
Eudyptes filholi	Yes	20000	60156.92	180925.56	230869.80	3.84
Aptenodytes patagonicus	Yes	19000	70151.67	167213.90	198392	2.83
Pygoscelis papua	Yes	17000	137119.22	158353.80	161620.11	1.18
Pygoscelis adeliae	Yes	15000	113114.94	156883.40	166240.67	1.47
Eudyptes c. chrysolophus/ Eudyptes c. schlegeli	Yes	15000	64332.57	149738.01	186514.43	2.90
Eudyptes moseleyi	No	NA	NA	158062.58	157782.11	NA
Eudyptes pachyrhynchus	No	NA	NA	188018.56	164099.70	NA
Pygoscelis antarctica	No	NA	NA	201910	187776.10	NA
Eudyptes chrysocome	No	NA	NA	243788.39	215299.52	NA

Supplementary Table 31. Fastsimcoal2 posterior distributions of N_e for *Eudyptes*, *Pygoscelis* and *Aptenodytes* at five different time points. 1 kya BP (Holocene), 15 kya BP (Last Glacial Maximum), 50 kya BP (Last Glacial), 140 kya BP (Penultimate Glaciation) and 1 mya BP. The fold increase/decrease is relative to the previous time period. Note that *Eudyptes chrysolophus chrysolophus*/*E. c. schlegeli* are combined in the same analysis. *E. robustus* and *E. moseleyi* were not analysed due to low sample size.

Taxon	Time point	mean	L_CI	H_CI	Increase/decrease	Fold increase/decrease
Eudyptes chrysolophus chrysolophus* and *E. c. schlegeli	1 mya BP	293,049	281,910	304,716	NA	NA
	140 kya BP	159,752	105,131	220,669	Decrease	1.83
	50 kya BP	608,175	89,918	1,373,423	Increase	3.81
	15 kya BP	354,214	40,165	1,049,173	Decrease	1.72
	1 kya BP	131,077	3,511	528,139	Decrease	2.7
Eudyptes filholi	1 mya BP	340,388	321,742	356,896	NA	NA
	140 kya BP	135,605	72,113	198,461	Decrease	2.51
	50 kya BP	470,056	51,169	1,258,839	Increase	3.47
	15 kya BP	376,825	25,302	1,196,708	Decrease	1.25
	1 kya BP	355,869	1,333	1,261,788	Decrease	1.06
Pygoscelis adeliae	1 mya BP	227,114	214,190	241,835	NA	NA
	140 kya BP	139,844	85,983	191,935	Decrease	1.62
	50 kya BP	654,742	72,222	1,414,491	Increase	4.68
	15 kya BP	178,694	39,978	336,410	Decrease	3.66
	1 kya BP	872,881	3,990	1,527,019	Increase	4.88
Pygoscelis papua	1 mya BP	372,061	349,171	392,927	NA	NA
	140 kya BP	240,595	45,385	521,688	Decrease	1.55
	50 kya BP	611,857	26,113	1,404,753	Increase	2.54
	15 kya BP	63,642	8,573	59,257	Decrease	9.61
	1 kya BP	779,884	1,020	1,511,310	Increase	12.25

Taxon	Time point	mean	L_CI	H_CI	Increase/decrease	Fold increase/decrease
Pygoscelis antarctica	1 mya BP	234,714	227,842	242,730	NA	NA
	140 kya BP	159,433	134,483	183,542	Decrease	6.61
	50 kya BP	1,053,820	350,636	1,668,917	Increase	1.88
	15 kya BP	560,050	209,782	1,030,993	Decrease	1.62
	1 kya BP	906,255	146,909	1,644,659	Increase	1.47
Aptenodytes patagonicus	1 mya BP	236,411	227,027	246,940	NA	NA
	140 kya BP	178,117	104,048	253,534	Decrease	1.33
	50 kya BP	608,506	84,604	1,379,334	Increase	3.42
	15 kya BP	153,463	36,521	224,063	Decrease	3.97
	1 kya BP	726,532	4,313	1,477,877	Increase	4.73
Aptenodytes forsteri	1 mya BP	323,326	307,442	338,944	NA	NA
	140 kya BP	105,285	57,316	152,471	Decrease	3.07
	50 kya BP	698,668	36,672	1,458,574	Increase	6.64
	15 kya BP	89,722	20,083	77,504	Decrease	7.79
	1 kya BP	916,546	1,659	1,545,508	Increase	10.22
Eudyptes chrysocome	1 mya BP	277,475	267,774	288,210	NA	NA
	140 kya BP	123,414	95,123	147,919	Decrease	2.25
	50 kya BP	1,054,393	222,419	1,731,583	Increase	8.54
	15 kya BP	1,061,584	269,028	1,755,434	Increase	1.01
	1 kya BP	998,008	302,083	1,665,025	Decrease	1.06
Eudyptes pachyrhynchus	1 mya BP	342,416	326,050	357,490	NA	NA
	140 kya BP	113,263	62,364	168,453	Decrease	3.02
	50 kya BP	895,338	80,813	1,673,800	Increase	7.9
	15 kya BP	990,180	22,011	1,643,468	Increase	1.11
	1 kya BP	1,006,783	331,987	1,698,321	Increase	1.02

Supplementary Table 32. Model likelihood values as inferred by Fastsimcoal2.

Taxon	Log MaxEstLhood	Log MaxObsLhood
Eudyptes chrysolophus chrysolophus and *E. c. schlegeli*	-12872.481	-12543.264
Eudyptes filholi	-11020.386	-10938.23
Pygoscelis adeliae	-5679.908	-5433.501
Pygoscelis papua	-5640.506	-5627.953
Pygoscelis antarctica	-16887.056	-16627.721
Aptenodytes patagonicus	-7376.743	-7143.136
Aptenodytes forsteri	-8122.035	-7798.126
Eudyptes chrysocome	-10435.277	-10426.278
Eudyptes pachyrhynchus	-5601.985	-5591.842

Supplementary Table 33. Tajima's D for *Eudyptes*, *Pygoscelis* and *Aptenodytes* penguins as inferred by δaδi. Negative values suggest a population expansion and positive values suggest a stable or negative N_e. *Note *Eudyptes pachyrhynchus* did not have a statistically significant P-value, even though it had a negative Tajima's D.

Taxon	Expansion	Tajima's D	P- value
Eudyptes chrysolophus chrysolophus and *E. c. schlegeli*	Yes	-0.77	<0.001
Eudyptes filholi	Yes	-0.21	<0.001
Pygoscelis adeliae	Yes	-0.77	<0.001
Pygoscelis papua	No	0.77	<0.001
Pygoscelis antarctica	Yes	-0.31	<0.001
Aptenodytes patagonicus	Yes	-1.15	<0.001
Aptenodytes forsteri	Yes	-0.76	<0.001
Eudyptes moseleyi	No	0.13	<0.001
Eudyptes chrysocome	No	0.73	<0.001
Eudyptes pachyrhynchus	Yes*	-0.03	0.135
Eudyptes robustus	No	0.16	<0.001

Supplementary Table 34. Summary of Multi-dice results. The median, mean and mode for the time of a synchronous expansion event (in years) is shown, and the proportion of co-expanding taxa.

	Median	Mean	Mode
Time of synchronous expansion event (years)	10,604	14,342	7,126
Proportion of co-expanding taxa	0.43	0.42	0.28

Chapter 4

Ancient DNA of crested penguins: Testing for temporal genetic shifts in the world's most diverse penguin clade

Publication details, contributions and acknowledgements

Parts of this chapter are published in:

Theresa L Cole[1,2], Nicolas J Rawlence[1], Nicolas Dussex[3,4], Ursula Ellenberg[5,6], David M Houston[7], Thomas Mattern[1], Colin M Miskelly[8], Kyle W Morrison[9], R Paul Scofield[10], Alan JD Tennyson[8], David R Thompson[11], Jamie R Wood[2] & Jonathan M Waters[1]. (2019). Ancient DNA of crested penguins: Testing for temporal genetic shifts in the world's most diverse penguin clade. *Molecular Phylogenetics and Evolution.* 131(1): 72 – 79.

Author affiliations:

[1]Department of Zoology, University of Otago, PO Box 56, Dunedin 9054, Otago, New Zealand. [2]Manaaki Whenua Landcare Research, PO Box 69040, Lincoln 7640, Canterbury, New Zealand. [3]Department of Bioinformatics and Genetics, Swedish Museum of Natural History, Box 50007, Stockholm 10405, Sweden. [4]Department of Anatomy, University of Otago, PO Box 56, Dunedin 9054, Otago, New Zealand. [5]Department of Ecology, Environment and Evolution, La Trobe University, Melbourne 3083, Victoria, Australia. [6]Global Penguin Society, University of Washington, Seattle 98195, United States of America. [7]Biodiversity, Department of Conservation, Private Bag 68908, Wellesley Street, Auckland 1141, Auckland, New Zealand. [8]Museum of New Zealand Te Papa Tongarewa, PO Box 467, Wellington 6140, Wellington, New Zealand. [9]964 Oriole Drive, Peterborough, Ontario K9H6K7, Canada. [10]Canterbury Museum, Rolleston Avenue, Christchurch 8001, Canterbury, New Zealand. [11]National Institute of Water and Atmospheric Research Ltd., Private Bag 14901, Kilbirnie 6241, Wellington, Wellington, New Zealand.

Author contributions:

Theresa L Cole (40%), Jonathan M Waters (30%) and Nicolas J Rawlence (30%) conceived the study. Nicolas J Rawlence (90%), Theresa L Cole (5%) and Lara D Shepherd (5%) sampled specimens from museums. David R Thompson (12%), Theresa L Cole (12%), Ursula Ellenberg (12%), Thomas Mattern (12%), David M Houston (12%), Kyle W Morrison (12%), Alan JD Tennyson (12%), Colin M Miskelly (12%) and Nicolas J Rawlence (4%) collected modern blood samples. Theresa L Cole (80%), Nicolas J Rawlence (15%) and Lara D Shepherd (5%) undertook laboratory work. Theresa L Cole (80%), Nicolas Dussex (10%), Nicolas J Rawlence (5%) and Jamie R Wood (5%) analysed the data. Theresa L Cole (40%), Jamie R Wood (30%) and Jonathan M Waters (30%) contributed sequencing and laboratory costs. All co-authors contributed to writing the manuscript.

Acknowledgements:

This work was supported by an Australasian Seabird Group grant and Manaaki Whenua Landcare Research. Theresa L Cole was supported by an Otago Postgraduate Scholarship. The authors thank Alastair Judkins (Kaikoura Ocean Research Institute) for providing a prehistoric penguin skull for

analysis. Leo Joseph, Robert Palmer, and Brian Gill assisted with arranging sampling from museum collections. Fieldwork would not have been possible without Hotte Mattern, Paul Sagar, Johanna Hiscock, Chris Rickard, Sarah Fraser, Kath Walker and Graeme Elliott. We are grateful to Tom Hart, Gary Miller, Paul Nolan, Petra Quillfeldt and Lara D Shepherd for providing blood samples that were used as the outgroups in *Supplementary Figure 24*, and to Barbara Wienecke for tracking down sample information. We thank several anonymous reviewers and Gustavo Cabbane for improving the manuscript. We consulted the Ngāi Tahu Research Consultation Committee about this project. The research was approved by the Otago University Animal Ethics Committee 61/2016 and conducted under NZ Department of Conservation authority numbers: OT-25557-DOA (Bruce Robertson), 32202-FAU, 35682-FAU, 37312-LND, 50437-DOA (David Thompson), 50436-FAU, 50464-DOA (Theresa Cole), and 38882-RES (Thomas Mattern and Philip Seddon).

Abstract

Human impacts have substantially reduced avian biodiversity in many parts of the world, particularly on isolated islands of the Pacific Ocean. The NZ archipelago, including its five sub-Antarctic island groups, holds breeding grounds for a third of the world's penguin species, including several representatives of the diverse crested penguin genus *Eudyptes*. While this species-rich genus has been little studied genetically, recent population estimates indicate that several *Eudyptes* taxa are experiencing contemporary demographic declines. Although *Eudyptes* are currently limited to the southern regions of the NZ archipelago, fossil and archaeological deposits suggest a wider distribution during prehistoric times, with breeding ranges perhaps extending to the North Island. Here, we analyse ancient, historic and modern DNA sequences to explore two hypotheses regarding the recent history of *Eudyptes* in NZ, testing for (1) human-driven extinction of *Eudyptes* lineages; and (2) reduced genetic diversity in surviving lineages. From 83 prehistoric bone samples, each tentatively identified as '*Eudyptes* spp.', we genetically identified six prehistoric penguin taxa from mainland NZ, including the extinct *E. warhami*. Moreover, our demographic analyses indicated that, while the breeding range of the *E. pachyrhynchus* may have contracted markedly over the last millennium, genetic DNA diversity within this lineage has remained relatively constant. This result contrasts with human-driven biodiversity reductions previously detected in several NZ coastal vertebrate taxa.

Introduction

Humans have had major impacts on island biota (Duncan *et al*., 2013; Wood *et al*., 2017). Hunting, habitat loss and the introduction of invasive predators accompanied human settlement and caused declines and extinctions of many island species (Spatz *et al*., 2017; Wood *et al*., 2017), particularly those endemic to the Pacific Islands (Steadman, 1989; Ceballos *et al*., 2015),

including NZ (Worthy & Holdaway, 2002). The unique evolutionary history of NZ's vertebrate biota has long been the subject of particular scientific intrigue (Fleming, 1979; Diamond, 1990). With prolonged geographic isolation, and in the absence of mammalian predators, many native avian taxa evolved distinctive ecological adaptations and lost anti-predator defences. NZ was the world's last major landmass to be settled by humans, with Polynesian arrival about 780 years ago (Wilmshurst *et al*., 2008; Wilmshurst *et al*., 2011; Anderson, 2017), followed 520 years later by European colonisation. Subsequently NZ's biota has experienced a well-documented decline that is clearly associated with human activities (Worthy, 1999; Worthy & Holdaway, 2002) over a period of relative climatic stability (Wanner *et al*., 2008; Rawlence *et al*., 2016; Waters *et al*., 2017). As a result of hunting (Rawlence *et al*., 2015a), rapid landscape modification (McWethy *et al*., 2010) and predation by introduced species that followed both colonisation events, a substantial proportion of NZ's endemic vertebrates became extinct (Tennyson & Martinson, 2007). Those species unable to fly, including many birds, were particularly susceptible to these rapid ecological changes. Since initial human settlement, 61 endemic birds have gone extinct (Tennyson & Martinson, 2007), 71 are now threatened with extinction, and 23 are classified as Nationally Critical (Robertson *et al*., 2016).

Recently, a number of studies combined ancient and contemporary DNA analyses with morphological analyses and radiocarbon dating, to explore the impact of human settlement on population size in a range of extant and extinct NZ animal taxa (see Cole & Wood, 2017b), including sea lions (*Phocarctos hookeri*; Collins *et al*., 2014a; Collins *et al*., 2014b; Rawlence *et al*., 2016), fur seals (*Arctocephalus* spp.; Salis *et al*., 2016), kiwi (*Apteryx* spp.; Shepherd & Lambert, 2008; Shepherd *et al*., 2012), swans (*Cygnus* spp.; Rawlence *et al*., 2017a), kākāpō (*Strigops habroptilus*; Bergner *et al*., 2016; Dussex *et al*., 2018), kea (*Nestor notabilis*; Dussex *et al*., 2015), huia (*Heteralocha acutirostris*; Dussex *et al*., 2019), kōkako (*Callaeas cinereus*; Dussex *et al*., 2019) South Island shags (*Leucocarbo* spp.; Rawlence *et al*., 2015b, 2017a) and penguins (*Megadyptes* and *Eudyptula* spp.; Boessenkool *et al*., 2008; Rawlence *et al*., 2015a; Grosser *et al*., 2015; Grosser *et al*., 2016). Several of these analyses have discovered cryptic extinctions and population declines in coastal and marine fauna following human settlement, and subsequent range expansions of congeneric species. For example, the presence of bones attributed to *Megadyptes antipodes antipodes* in natural and archaeological bone deposits around the NZ coast led to the assumption that this species had been present on the mainland before human settlement. By combining ancient and contemporary DNA with morphological analyses, Boessenkool *et al*. (2008) overturned this hypothesis, revealing that a previously unknown *Megadyptes* taxon, *M. a. waitaha* had inhabited the NZ coast prior to human arrival. Following human settlement *M. a. waitaha* was likely hunted to extinction, and was

subsequently rapidly replaced by an expanded population of *M. a. antipodes* from the sub-Antarctic (Rawlence *et al*., 2015a). Moreover, Grosser *et al*. (2015) and Grosser *et al*. (2017) identified that two species of *Eudyptula* now breed around the NZ coast; the endemic *Eudyptula minor* and the Australian *E. novaehollandiae*. Grosser *et al*. (2016) demonstrated that the presence of *E. novaehollandiae* in NZ represents a cryptic colonisation event from Australia, that probably followed the localised extirpation of the Otago *E. minor* lineage. However, little is known about whether similar extinction/colonisation events occurred within prehistoric crested penguins (*Eudyptes* spp.) in NZ, even though skeletal remains belonging to this genus have been collected from natural Pleistocene and Holocene subfossil and midden deposits around NZ (Scarlett, 1979; Leach, 1979; Millener, 1990; Worthy, 1997).

New Zealand provides important breeding grounds for half of the world's *Eudyptes* penguins. Endemic *Eudyptes* include *E. sclateri*, which breeds on the Antipodes and Bounty Island groups, *E. robustus* which breeds on The Snares and Western Chain islets; and *E. pachyrhynchus*, which breeds in dense forest along the coast of southern Westland and Fiordland in the South Island, Stewart Island and smaller offshore islands. Several studies have demonstrated major declines in *Eudyptes* during the late 20th century (Crawford *et al*., 2009; Crawford *et al*., 2010; Davis, 2013; Morrison *et al*., 2015; Otley *et al*., 2018). Taylor (2000) and Otley *et al*. (2018) suggest *E. pachyrhynchus* populations have declined significantly during this time. Surveys of *E. sclateri* suggested that populations were stable until 2011, but subsequently declined by 19 percent (Hiscock & Chilvers., 2014). In contrast, Warham (1974a), Mattern (2013) and Hiscock and Chilvers (2016) suggested that *E. robustus* populations have remained stable. By contrast, the extent to which any such recent demographic shifts might have impacted the genetic composition of *Eudyptes* populations remains unclear.

Here, we investigate the prehistoric distribution and population history of *Eudyptes* species in NZ by comparing genetic information derived from prehistoric bones, historical museum skins, and blood from extant populations. Specifically, we test two hypotheses regarding the effects of human colonisation on *Eudyptes* diversity: (1) human-driven extinction of *Eudyptes* lineages; and (2) temporal declines in genetic diversity within extant *Eudyptes* lineages.

Materials and Methods

Source of specimens

Sub-fossil and archaeological bones (n = 83) tentatively identified as *Eudyptes* spp. based on size, and historical museum skins (n = 27, in addition to 21 sequences obtained from GenBank;

Supplementary Tables 35 and *36*) were sourced from collections at the NMNZ, Canterbury Museum, Auckland War Memorial Museum and the Australian National Wildlife Collection (ANWC). To ensure independence of prehistoric samples, only common elements of the left or right orientation were sampled from an individual deposit, or else bones were sampled from different stratigraphic units within a site. In addition, whole blood or tissue (n = 87) was collected from wild individuals across the breeding range of *E. pachyrhynchus*, *E. robustus* and *E. sclateri* (except the Bounty Islands) (*Supplementary Table 35*). Our final sample size consisted of 268 specimens (including additional specimens obtained from GenBank) collected across NZ and the sub-Antarctic region, over a temporal scale encompassing the time period from initial human arrival to the present (i.e. the last 780 years before present). Ages were determined based on stratigraphic context (archaeological details and radiocarbon dates provided in *Supplementary Table 35*).

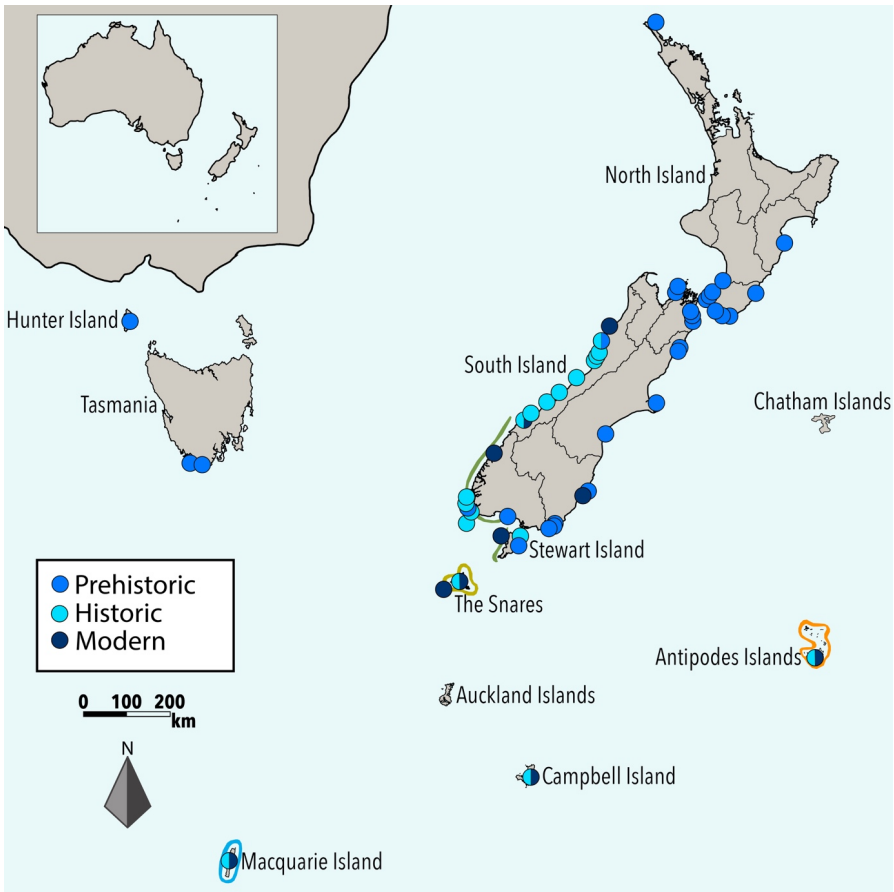


Figure 21. Sampling locations for prehistoric, historic and modern NZ penguin samples. Prehistoric indicates prehistoric sub-fossil and archaeological bones, Historic indicates historical museum skins and Modern indicates contemporary blood. Historical museum skins that do not have locality collection details associated with them (*Supplementary Tables 35* and *36*) are not shown on the map. Modern breeding sites of *Eudyptes* penguins are shown; *E. pachyrhynchus* (green), *E. robustus* (yellow), *E. sclateri* (orange) and the Macquarie Island-endemic *E. chrysolophus schlegeli* (blue). Samples from sub-Antarctic islands outside the Australian and NZ region are not shown, but can be referred to in *Supplementary Table 35*. Historical Macquarie Island samples were obtained from Cole *et al*. (2018; *Chapter 5*). Note *E. c. schlegeli* breeds on Macquarie Island, which is outside of the NZ sub-Antarctic region.

DNA extraction, amplification and sequencing

DNA was extracted following several techniques, optimised for the sample type used. aDNA extraction and PCR set-up for sub-fossil bones, archaeological bones and historical museum skins were carried out in two purpose-built aDNA laboratories that were physically isolated from other molecular laboratories, following rigorous aDNA protocols (Cooper & Poinar, 2000). DNA extractions were mostly performed at the Department of Zoology, University of Otago, Dunedin following Rohland *et al*. (2010) for sub-fossil and archaeological material and Rawlence *et al*. (2015a) for historical museum skins. Three prehistoric specimens from Hunter Island, Tasmania, were extracted at MWLR, Lincoln, following Thomson *et al*. (2014), as outlined in Cole *et al*. (2018; *Chapter 5*). DNA from blood was extracted following the Qiagen DNeasy® Tissue Extraction Kit (Qiagen Inc., Chatsworth, California, USA) at MWLR, Lincoln.

To verify and assign the historical museum skins and sub-fossil/archaeological bones to a species we followed the methods described by Cole *et al*. (2018; *Chapter 5*) to amplify up to 500 bp of the COI gene, using four overlapping primer pairs, appropriate for aDNA. To supplement the dataset, we sequenced COI from two samples of *Megadyptes antipodes waitaha* and three of *M. a. antipodes,* which have previously been differentiated based on their mitochondrial CR sequences. As the COI markers could only amplify one to two fragments for *Megadyptes* ssp., we assigned samples that were identified as *Megadyptes* ssp. to species level using the CR primers of Boessenkool *et al*. (2008). To test the demographic history of *Eudyptes* spp., we also designed *Eudyptes* specific primers (Eud_CR2F 5'GGGTTTAGTCCTCTCTATTGGG and Eud_CR1R 5'GCTTACTGCTGTGTTTCAGGGA) to amplify a 131 bp CR fragment that contains 23 variable sites, 16 of which are parsimoniously informative, within *Eudyptes* (*E. pachyrhynchus*: 12 variable/6 informative; *E. robustus*: 7 variable/3 informative; *E. sclateri*: 10 variable/7 informative). PCRs (12.5 µL each) were performed on sub-fossil/archaeological and historical skin material using 2 mg/mL BSA (Sigma), 1 X PCR buffer, 2 mM $MgSO_4$, 80 µM dNTPs, 0.4 µM of each primer, 0.625 U of HiFi Platinum Taq (Invitrogen) and 1 µL DNA on a BIO-RAD MyCycler thermal cycler with an initial denaturation of 94°C for 3 min; followed by 55 cycles of 94°C for 30 s, 55°C for 30 s and 68°C for 45 s; and a final extension at 68°C for 10 min. PCRs for amplifying the CR from blood samples were performed in 1 X PCR buffer, 250 µM dNTPs, 1.25 U of i-Taq (iNtRON), 0.25 µM of each primer, and 1 µL DNA, following the same PCR conditions while lowering the number of cycles to 42 and raising the annealing temperature to 58°C. PCR products were purified using SPRIselect (Beckman Coulter, Inc.,

Indianapolis, IN, USA) or exoSAP-IT™ (ThermoFisher Scientific), for sub-fossil/archaeological/historical museum skins and blood samples respectively, and sequenced at the MWLR sequencing facility, Auckland, on an Applied Biosystems 3500xL Genetic Analyzer.

Contiguous sequences of both COI and CR were constructed and aligned using Geneious 8.1.8 (Biomatters) from high quality bidirectional sequence reads and checked manually. Due to post-mortem DNA damage, when inconsistency between sequences from a given individual was observed (e.g. G to A and C to T transitions), additional PCR and bi-directional sequencing was conducted, and a majority rule consensus was applied (Brotherton *et al*., 2007).

Phylogenetic analysis

To verify and assign the samples to a species, we created two Bayesian maximum-credibility phylogenies using COI, and one Bayesian maximum-credibility phylogeny using the *Megadyptes* specific CR. As is typical with aDNA, our data contained some partial sections, e.g. missing sequence data. We therefore constructed a first phylogeny which contained all COI data, and a second phylogeny that contained only samples with three to four of the four overlapping fragments. In addition, we included one sequence representative of all extant Sphenisciform species, and *Diomedia exulans* as an outgroup from GenBank. For the *Megadyptes* phylogeny, we obtained an additional 15 and 17 *M. antipodes waitaha* and *M. a. antipodes* sequences, and one *E. chrysocome* sample as an outgroup from GenBank (*Supplementary Table 36*). In addition, we created a CR phylogeny to test for phylogeographic structure among *E. pachyrhynchus* samples. All phylogenetic trees were created using BEAST 2.4.7 (Bouckaert *et al*., 2014), with a relaxed log-normal clock for COI and a strict clock for CR, with a Yule model of speciation for 100 million MCMC generations, sampling tree parameters every 1000 generations with a burn-in of 10%. Each tree was replicated in triplicate and combined using Log Combiner. We implemented the AIC in JModelTest2 (Darriba *et al*., 2012) to determine the most appropriate model (Jukes Cantor for all genetic markers).

Temporal haplotype network

We created temporal haplotype networks using CR sequences for *E. pachyrhynchus, E. sclateri* and *E. robustus* by employing the TempNet R script (Prost & Anderson, 2011) to visualise haplotypic changes over a temporal scale for sub-fossil/archaeological bones, historical museum skins and blood for each time period. We amplified 111 bp of the CR from a *Eudyptes* taxon which fell separate to *E. pachyrhynchus*, *E. sclateri* and *E. robustus* (see also Cole *et al*.

2019b; *Chapter 2*), so we created additional networks that included a representative of all *Eudyptes* species, except *E. moseleyi* because no sequence was available.

Estimating temporal fluctuations of N_e

We reconstructed the demographic history of *E. pachyrhynchus* by creating a Bayesian Skyline Plot (BSP) (Drummond *et al*., 2005) in BEAST 2.4.7. Tip dates, in mean calibrated years before present (where present is 2016) were determined for all samples based on (1) dates of associated biological material from the same archaeological sites, excavation squares or stratigraphic units; (2) associated faunal/cultural assemblages; and (3) collection date for historical skins (see *Supplementary Table 35*). We removed the sample AD357 (from North Cape, Northland) from the analysis because we were not confident about its calibrated age (between 6 Ka BCE to 1450 CE) and collection locality (an outlier to all other locations). We found no phylogeographic structure within *E. pachyrhynchus* (see *Figure 22*) so the assumption of population panmixia was not violated (Drummond *et al*., 2005). We estimated the timing of demographic events of *E. pachyrhynchus* using a constant, exponential increase, Bayesian Skyline coalescent prior in BEAST 2.4.7 using the Jukes Cantor model of nucleotide substitution and a strict molecular clock model of evolution. Each run was conducted three times, consisting of 50 million MCMC generations, sampling tree parameters every 1000 generations,with a pre-burn-in of 10%. An extended Bayesian Skyline coalescent prior was also attempted, extending the MCMC generations to 100 million, 150 million, 300 million and 500 million, however for those analyses the ESS values did not converge. A three-way comparison of each model produced high substitution rates (>2 substitutions/site/Ma). The Bayes Factor comparison indicated that the most likely model was the BSP, and the second most likely model was a constant population size. As the model likelihood values were only within two of each other, which may indicate low information content of our data (Ho *et al*., 2011), we chose the simpler model of constant population size as recommended by Ho *et al*. (2011). Nevertheless, we constructed the BSP using Tracer v1.6.0. The mean substitution rates for the BSP were 2.13 substitutions/site/Ma (95% CIs: 1.14 – 4.94 substitutions/site/Ma), and the constant population size was 2.69 substitutions/site/Ma (95% CIs: 2.71 – 5.89 substitutions/site/Ma). This is a similar rate to other seabirds, e.g. the Otago shag (*L. chalconotus*) has 3.20 substitutions/site/Ma and the Foveaux shag (*L. stewarti*) has 1.73 substitutions/site/Ma (Rawlence *et al*., 2015b), however is faster than *Pygoscelis adeliae* which has 0.29 – 0.88 substitutions/site/Ma (Millar *et al*., 2008).

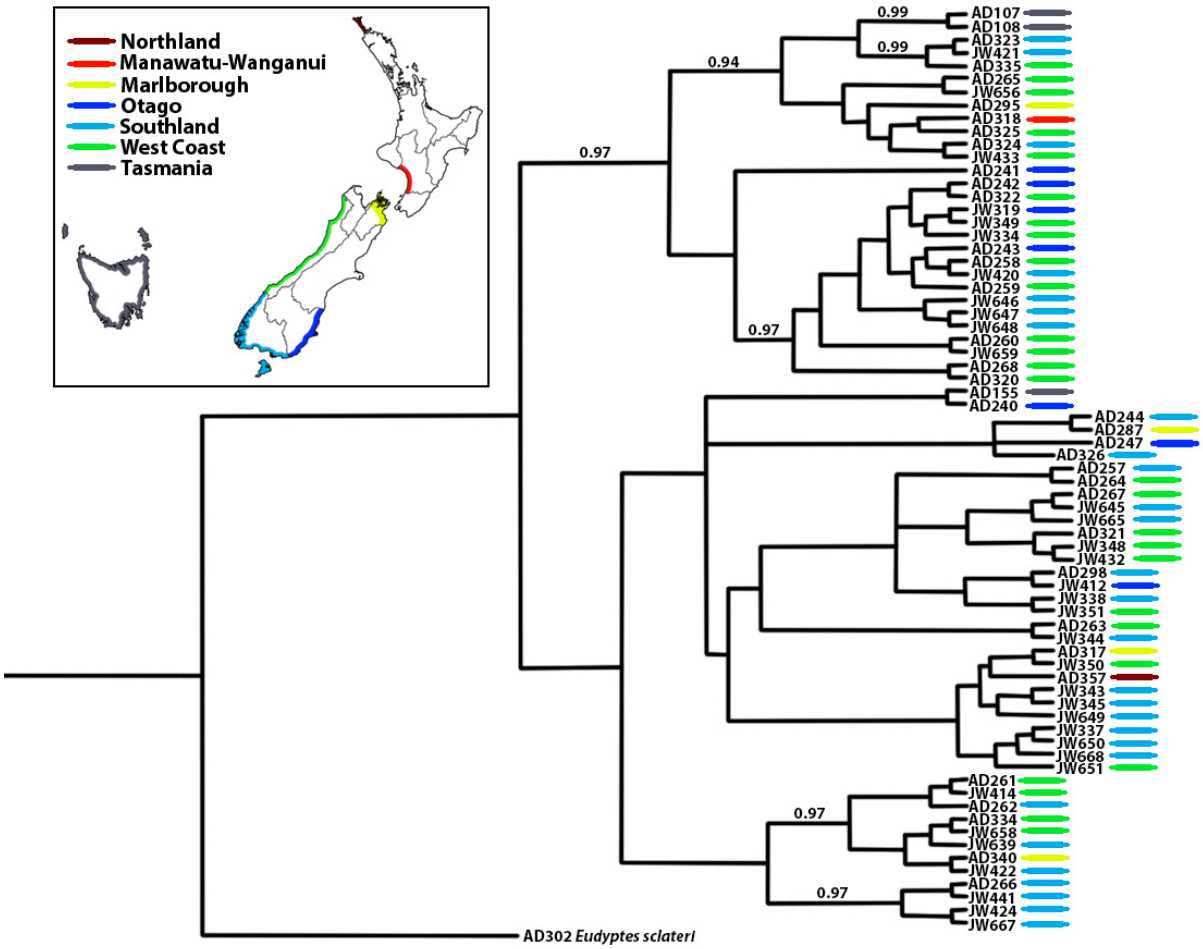


Figure 22. Bayesian phylogenetic tree of *Eudyptes pachyrhynchus* based on the CR. Coloured ovals represent the regional district that the sample originated from. Posterior probabilities >0.91 are shown.

Given the shallow time-depth of our analyses (1151 cal years BP), we assessed whether our data contained sufficient temporal genetic structure to allow reliable estimates of substitution rates by randomising the dates across samples in 20 independent replicates and repeating the constant population size model analysis, as suggested by Ho *et al*. (2011), for low information content aDNA datasets. The 95% CIs of estimated substitution rates were visualised in Tracer. If the original mean rate-estimate was contained within the 95% CI of the randomised data sets, there is likely insufficient temporal structure for reliable rate estimation (Ho *et al*., 2011). A more conservative approach argues that substitution rates are unreliable if the 95% CI of the original rate-estimate and randomised data sets overlap (Duchene *et al*., 2015). Caution therefore needs to be taken over the reliability of our mean substitution rate estimate despite being similar to those of other seabirds (see also Rawlence *et al*., 2015b) given the wide CIs and that the temporal CR dataset failed both randomisation tests, i.e. over half the 95% CIs overlapped with the mean empirical rate estimate. This caveat reflects the fact that short mtDNA CR fragments provide only limited information for substitution rate estimation (Ho *et al*., 2007; Ho *et al*., 2011).

Testing demographic scenarios

We used an ABC approach (Beaumont *et al*., 2002) implemented in DIYABC 2.1.0 (Cornuet *et al*., 2014) to test for population bottlenecks and to estimate values of key demographic parameters. Specifically, we designed models to test for the potential demographic declines in *Eudyptes pachyrhynchus* associated with LGM climate change (see also *Chapter 3*) and human arrival in NZ. We assumed a generation time of five years based on the age of first breeding (Otley *et al*., 2017). We did not analyse *E. robustus* or *E. sclateri* due to low sample size. Four models were tested: (1) a model of constant population size since the Holocene with N_e uniformly distributed through time (>3000 generations; >50 Ka cal years BC); (2) a model of Holocene decline (2000 – 3000 generations; 10 – 15 Ka cal years BP); (3) a model of human-induced decline (2 – 160 generations, 10 – 800 cal years BP); and (4) a model of Holocene (2000 – 3000 generations; 10 – 15 Ka cal years BP) and human-induced (2 – 160 generations, 10 – 800 cal years BP) declines (*Supplementary Figure 25*). Recent surveys suggest that the current population estimates are at least 5500 – 7000 mature *E. pachyrhynchus* individuals (BirdLife International, 2016) and there is no prior information on pre-human population size in the species available. We therefore used wide priors (*Supplementary Table 37*) for pre-LGM, pre-human and modern female effective population size (N_{ef}). For this reason, we set up as condition of the ABC analysis the pre-LGM and pre-human N_{ef} to be larger than the modern N_{ef} in all the scenarios tested in order to simulate a population bottleneck (constraints applied to N_e: modern N_{ef} < pre-LGM N_{ef} and modern N_{ef} < pre-human N_{ef}). For each model, 1 million datasets were simulated with the defined demographic and marker parameters (*Supplementary Table 38*). Samples were grouped into three temporal samples corresponding to a modern (>50 cal years BP), a historical (50 – 100 cal years BP) and a prehistorical period (500 – 1000 cal years BP). In our models, the collection times of these samples were thus set at 0, 20 and 200 generations. We used a Jukes Cantor substitution model with a mean rate uniformly distributed between 1.00×10^{-8} and 1.00×10^{-7} substitutions/generations. For summary statistics, we used the number of haplotypes, number of segregating sites, mean of pairwise differences, and Tajima's D.

For model selection and parameter estimation, normalized Euclidean distances were calculated between the observed dataset and each of the simulated datasets using the local linear regression method of Beaumont *et al*. (2002). The posterior probabilities of each model were estimated by taking a logistic regression approach on the closest 1% of data sets to the observed data, providing both point estimated and 95% CIs (Cornuet *et al*., 2010; Fagundes *et al*., 2007). The posterior parameter distribution was estimated by retaining the 10,000 datasets for the preferred

model (1%) with the smallest Euclidean distances to the observed dataset (*Supplementary Table 37*). To check model performance, we first empirically evaluated the power of the model to discriminate among models, to ensure confidence in our model choice. We simulated pseudo-observed data sets with parameters drawn from the posterior parameter distribution of the preferred model and positioned the summary statistics of the observed data in the summary statistic distribution of these pseudo-observed datasets. The model was then considered suitable if the observed data summary statistics were included in the CIs drawn from pseudo-observed data. We then calculated statistical measures of performance and Type 1 and Type 2 error rates as a means of model checking (Cornuet *et al*., 2010; Excoffier *et al*., 2005).

Results

We successfully amplified COI and/or CR fragments from all blood samples and historical museum skins. Specifically, from a total of 83 sub-fossil and archaeological bones we amplified partial COI fragments from 41 samples, *Eudyptes* specific CR from 22 samples, and *Megadyptes* specific CR from nine samples. While the majority of samples (n = 32, including those obtained from GenBank; see *Supplementary Tables 35* and *36*) amplified for both mtDNA regions, a subset of sub-fossil and archaeological specimens failed to amplify for one or the other region (13 samples amplified COI but not CR, and four samples amplified CR but not COI; details in *Supplementary Table 35*). In total, the data obtained in our study captured a total of 58 new COI sequences, 145 new *Eudyptes* CR sequences and nine new *Megadyptes* CR sequences (*Supplementary Table 35*). All sequences were deposited on GenBank (details provided in *Supplementary Table 35*).

Phylogenetic relationships

We amplified COI from 41 sub-fossil and archaeological bones that had previously been tentatively identified as *Eudyptes* or *Megadyptes* based on morphological assessment, including five known *Megadyptes* penguins. Almost a third of the COI sequences (n = 16) fell within *Eudyptes*' sister genus *Megadyptes* (*Figure 23*), although the posterior probabilities for these phylogenetic assignments using COI was sometimes low, possibly due to some specimens having limited amplification success (average 175 bp) and/or the relatively small genetic divergence values among some of these lineages. Use of *Megadyptes* specific CR primers similarly identified nine of these samples as representing the extinct *M. a. waitaha* (*Figure 23*).

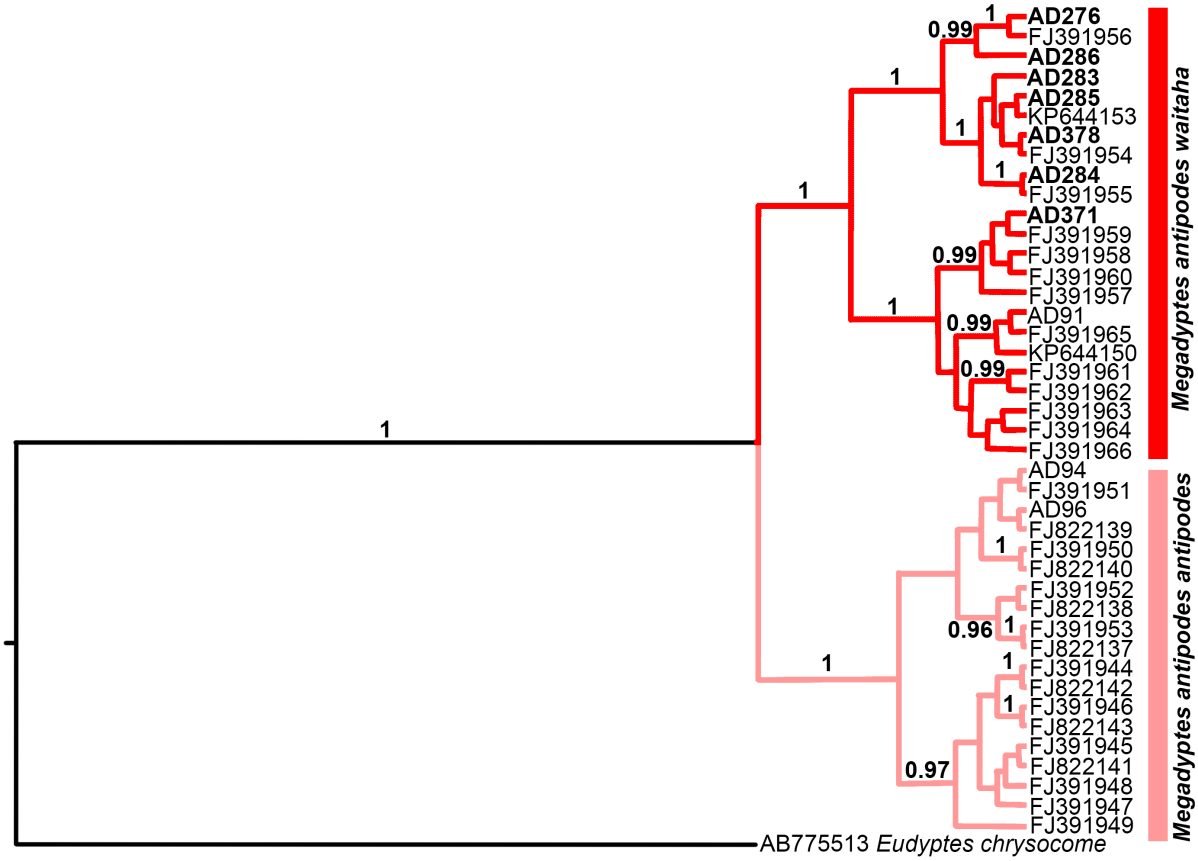


Figure 23. Bayesian phylogenetic tree of *Megadyptes* penguins with aDNA CR sequences from NZ bone samples. Seven *Megadyptes* prehistoric sub-fossil and archaeological bone samples were identified (in bold; see also *Supplementary Figure 20*). Posterior probabilities for the main *M. a. antipodes* and *M. a. waitaha* clades are shown.

We could not amplify CR for the remaining six samples so we only tentatively refer to them as *Megadyptes* spp. When we removed the short sequences, including all but two attributed to *Megadyptes*, the COI sequences successfully differentiated every *Eudyptes* species (except *E. schlegeli/E. chrysolophus* complex), including the occasionally merged *E. pachyrhynchus*/*E. robustus* (*Figure 24*). Of the remaining 25 samples, one is potentially attributable to *E. sclateri, E. schlegeli*/*E. chrysolophus,* one is either *E. chrysolophus chrysolophus*/*E. c. schlegeli*, three are *E. sclateri*, two are *E. robustus*, and 11 are *E. pachyrhynchus* (*Figure 24*; see *Supplementary Figure 22* for an alternative placement of AD258). Seven samples formed a separate clade within *Eudyptes,* sister *E. sclateri*, yet contains no extant samples (*Figure 24* and *Supplementary Figure 22*). While we refer to this taxon as *Eudyptes* clade X here (see Cole *et al*. 2019a), subsequent analyses of these samples in Cole *et al*. (2019b; *Chapter 2*) demonstrates that they are *E. warhami*.

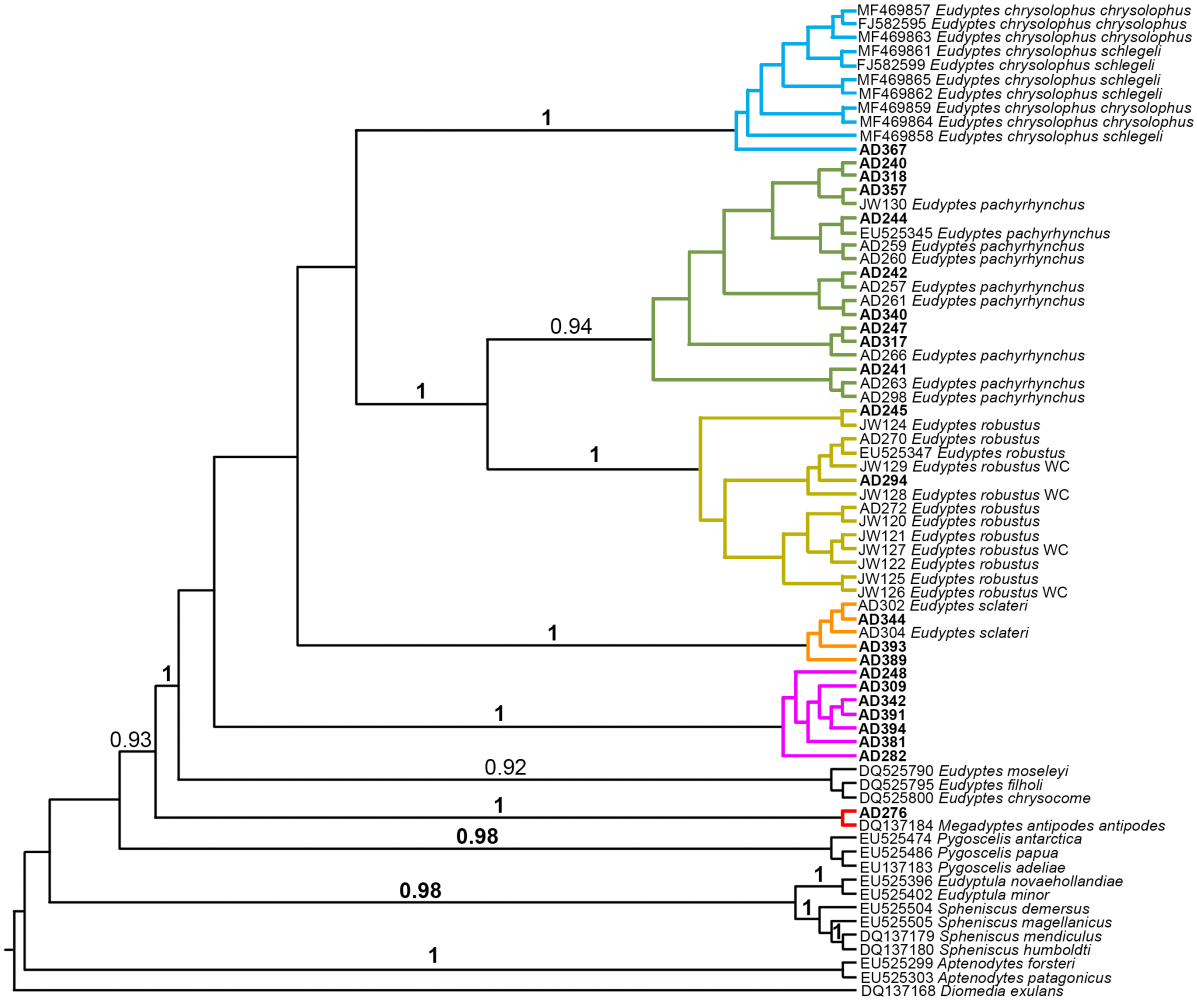


Figure 24. Bayesian phylogenetic tree of penguins, with high-quality aDNA COI sequences from NZ bone samples. *Eudyptes robustus* samples from the Western Chain are indicated with WC. AD numbers with the species name indicated were from historical museum skins, whereas samples without the species names (in bold) were prehistoric sub-fossil or archaeological bone samples. The sample size has been reduced from *Supplementary Figure 20* to include only samples with three to four of four overlapping COI fragments.

Population declines and range contractions

We amplified CR in 31 *E. robustus* individuals (*n* = 21 modern, *n* = seven historical, *n* = three prehistoric), 30 *E. sclateri* individuals (*n* = 18 modern, *n* = 10 historical, *n* = two prehistoric), 72 *E. pachyrhynchus* individuals (*n* = 34 modern, *n* = 23 historical and *n* = 15 prehistoric) and one sample of *E. warhami cf. Eudyptes* clade X (111 bp). The Snares and Western Chain *E. robustus* populations which breed six weeks apart were genetically indistinguishable from one another, supporting additional analyses of SNPs (see *Chapter 3*). Moreover, our data revealed only six haplotypes in *E. robustus,* but indicate there may have been no loss of genetic diversity in this species through time. Our data revealed only 10 haplotypes in *E. sclateri,* and there was also no loss of genetic diversity in this species through time. Of the 12 haplotypes that were observed in *E. pachyrhynchus*, the temporal network (*Figure 25* and *Supplementary Figures 23 – 24*) showed only a slight downward trend of genetic diversity over the last 1151 years

(sub-fossil/archaeological bones had nine haplotypes, historical museum specimens had six haplotypes and contemporary blood had six haplotypes). We did not find any phylogeographic structure when comparing CR sequences across *E. pachyrhynchus* specimens (*Figure 22*).

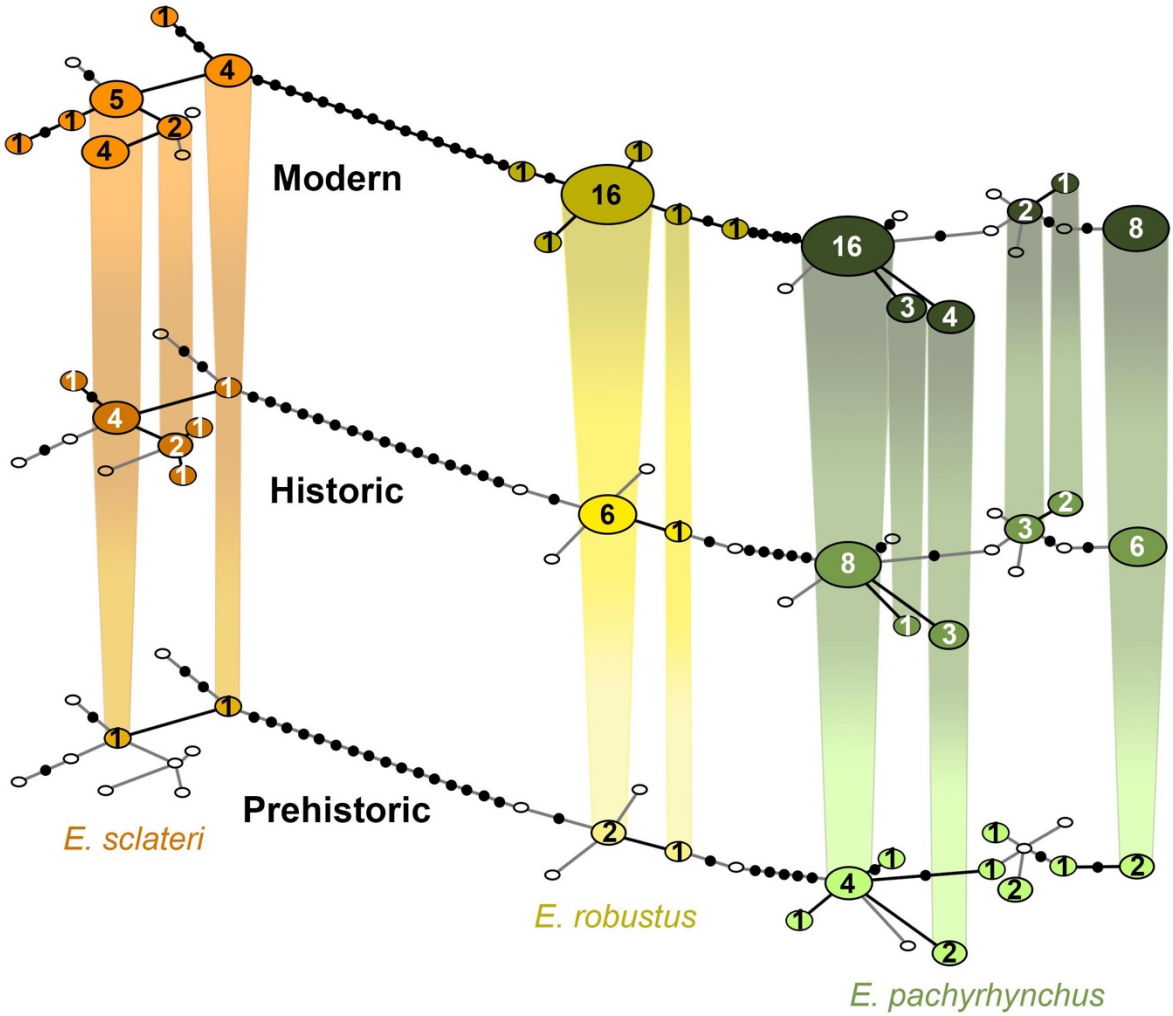


Figure 25. Temporal haplotype network of *Eudyptes sclateri*, *E. robustus* and *E. pachyrhynchus*. Coloured circles represent different haplotypes, with the size of the circle and number within the circle corresponding to the number of samples that had that haplotype. Small white circles represent those haplotypes that are present in one or two time periods, but are absent in the time period displayed. Black circles represent haplotypes that were absent from the dataset. Each layer represents a different time period, based on the sample type used (Prehistoric is sub-fossil and archaeological bone, Historic is historical museum skins and Modern is contemporary blood). The coloured cylinders connect those haplotypes that are present through multiple time periods. Green circles are *E. pachyrhynchus*, yellow circles are *E. robustus* and orange circles are *E. sclateri*.

When testing for competing demographic scenarios for *E. pachyrhynchus*, a model of constant population size over the last millennium was supported with a posterior probability of 67 percent (95% CI: 0.66 – 0.69) using the logistic regression approach (*Supplementary Tables 37* and *38*). This finding is consistent with the BSP (*Figure 26*). By contrast, alternative models of post-LGM, human and/or post-LGM and human-induced declines obtained a posterior probability of 10%, 13% and 10%, respectively. The observed data summary statistics for the

preferred model were included in the CI's drawn from pseudo-observed data (*Supplementary Figures 25* and *26*). Type 1 and Type 2 errors were 0.053 and 0.14, respectively for the preferred model of constant population size.

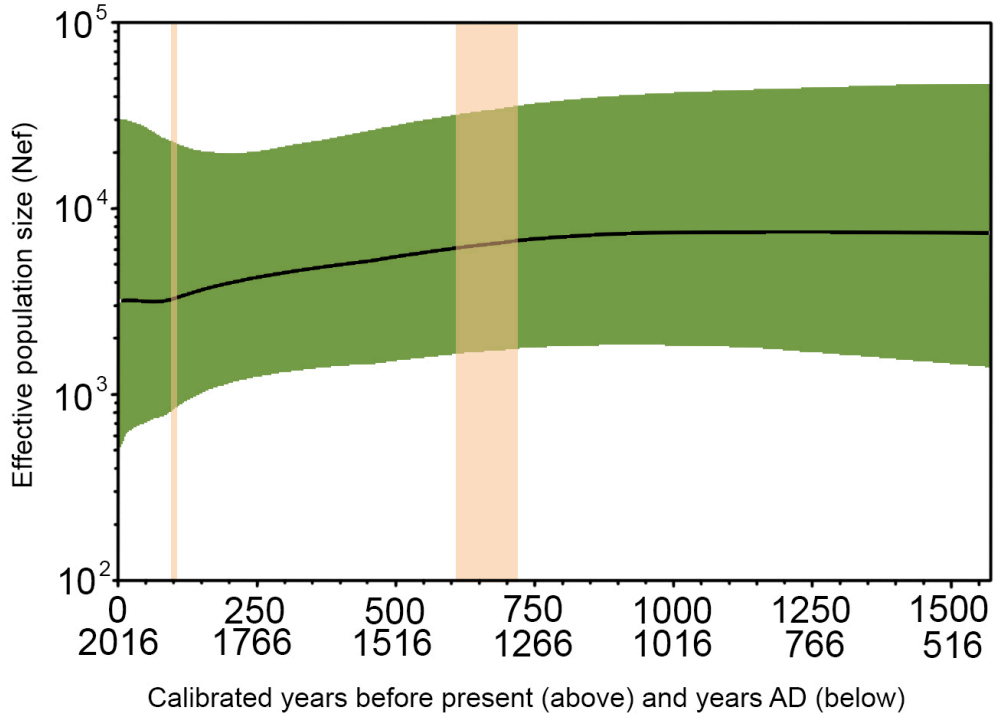


Figure 26. Bayesian Skyline Plot of of historical N_e for *Eudyptes pachyrhynchus*. The solid black line represenst the median estimate of N_e, while the shaded green areas represent the 95% CIs. The period from Polynesian colonization of NZ 1280 – 1450 CE during which NZ's megafauna was hunted to extinction (thick orange bar), and European arrival 1769 AD (thin orange bar) are shown.

Discussion

The present study presents genetic evidence for the formerly widespread distribution of *Eudyptes pachyrhynchus* throughout mainland NZ, which is consistent with previous morphological studies (Worthy, 1997). This is in contrast to our demographic models, which detected little change in population size, despite the apparent range reductions. Additionally, the detection of the extinct *E. warhami cf. Eudyptes* clade X in prehistoric NZ South and North Island archaeological and subfossil deposits raises further questions about the recent evolutionary history of this clade (also discussed in Cole *et al*. 2019b; *Chapter 2*).

Prehistoric penguin fauna

Our study genetically confirms the presence six large *Eudyptes* and *Megadyptes* penguin taxa (*Figure 27*) from mainland NZ's prehistoric record (see Worthy, 1997), a result which contrasts with the presence of just two contemporary breeding mainland representatives of these genera. The novel finding of a previously unknown and presumed extinct clade of *Eudyptes* (*Eudyptes*

warhami cf. Eudyptes clade X, represented by seven archaeological and subfossil samples) provides intriguing new evidence of a cryptic coastal vertebrate extinction in the NZ region. This potentially adds to the findings of recent aDNA studies (e.g. *Megadyptes, Phocarctos* and *Leucocarbo* extinctions; Boessenkool *et al*., 2008; Collins *et al*., 2014a; Rawlence *et al*., 2015a; Rawlence *et al*., 2015b; Rawlence *et al*., 2016; Rawlence *et al*., 2017b; Rawlence *et al*., 2018). It is unclear, however, whether *E. warhami cf. Eudyptes* clade X also represented a mainland breeding colony e.g. in the northern South Island or the lower North Island, or whether they were primarily vagrants from the Chatham Islands (see Cole *et al*., 2019b; *Chapter 2*). Notably, vagrant *Eudyptes* are not uncommon to NZ (Robertson, 2017), and we therefore also suspect that the relatively 'rare' mainland records of *E. warhami cf. Eudyptes* clade X, as well as the *E. sclateri*, *E. robustus,* and *E. c. schlegeli*/*E. c. chrysolophus* specimens, and the single *E. pachyrhynchus* from Northland, likely represent vagrants.

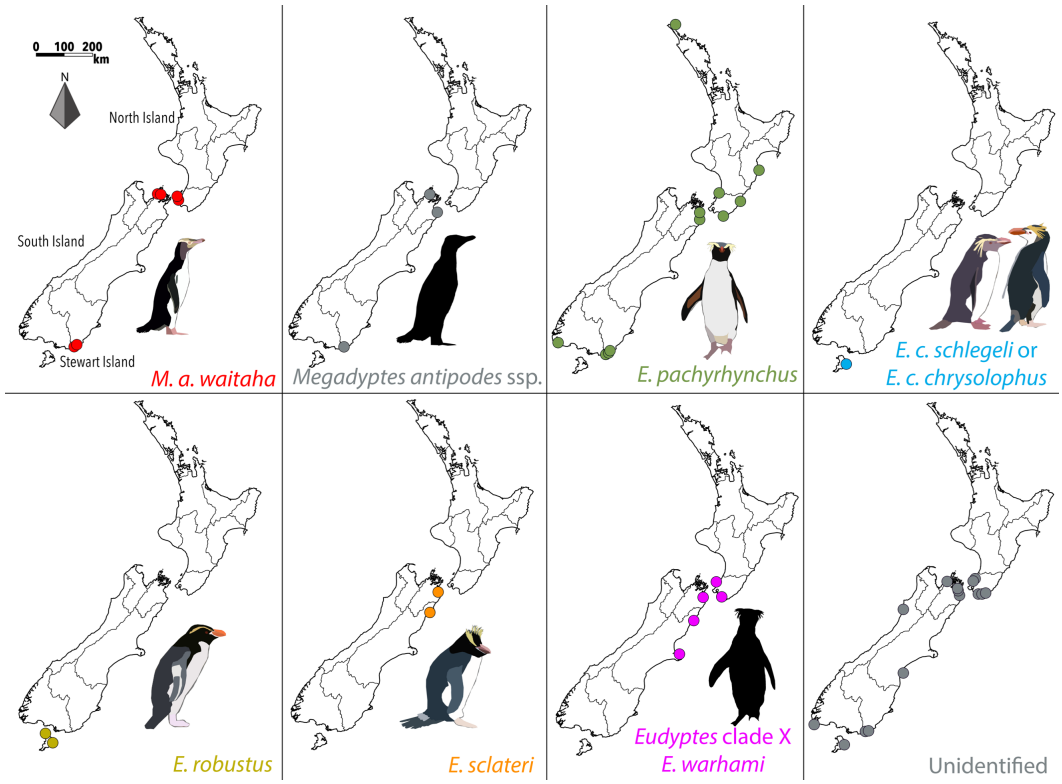


Figure 27. Penguin taxa identified from NZ bone samples. The map labelled 'Unidentified' indicates those samples that could not be amplified. *Megadyptes* samples in grey could not be identified to species level, so could be *M. a. waitaha*, *M. a. antipodes* or even *M. a. richdalei* (see Cole *et al*., 2019b; *Chapter 2*). *Megadyptes* identification is based on the CR.

No loss of genetic diversity despite range contractions

Based on our DIYABC analysis, our data supports a model of constant N_e in *E. pachyrhynchus* prior to human arrival in NZ, which suggests that the population has remained relatively stable. While the power of this demographic inference may be limited by the small sample size, however, based on our analyses it seems unlikely that hunting by early Polynesian settlers had

detectable species-wide effects on genetic diversity. These findings strongly contrast with recent studies on NZ coastal taxa such as the Otago shag (*Leucocarbo chalconotus*; Rawlence *et al*., 2015b) and NZ fur seals (*Arctocephalus forsteri*; Salis *et al*., 2016), which both experienced geographic and genetic declines. Importantly, in cases where populations are genetically well connected, range contractions do not necessarily result in substantially reduced N_e and genetic diversity, as seen in Matocq and Villablanca (2001), Hoffman and Blouin (2004) and Dussex *et al*. (2015).

Based on these findings, we hypothesise that the core population of *E. pachyrhynchus* has remained stable in the remote southern and southwestern regions of the NZ South Island over recent centuries. Similarly, Rawlence *et al*. (2015b) and Salis *et al*. (2016) demonstrated that following extirpation of northern lineages, endemic genetic lineages of *Leucocarbo* shags and *Arctocephalus* seals have persisted in southern NZ throughout human settlement. The relatively low human pressure in this remote southern region (Waters *et al*., 2017) may explain why *E. pachyrhynchus* has persisted in the southern mainland.

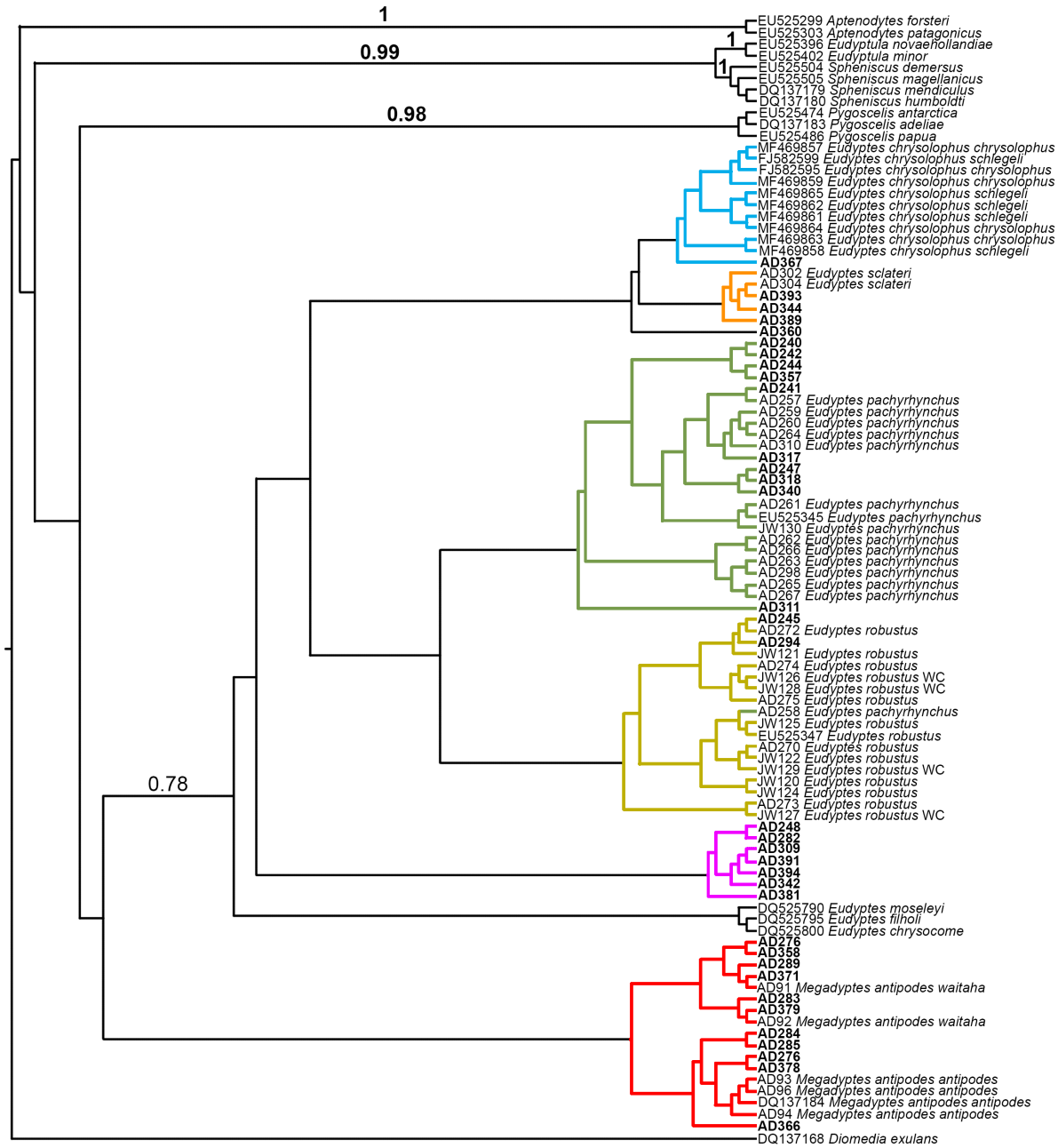
While our genetic data suggest that *E. pachyrhynchus* experienced anthropogenic range reductions in mainland NZ, alternative possibilities exist. The prehistoric presence of *Eudyptes* beyond its current range could instead represent occasional vagrant moulting individuals, rather than northern breeding colonies (Keogan *et al*., 2018). Moreover, Cole *et al.* (2018; *Chapter 5*) suggested there may be a bias towards large vagrant bird species in midden deposits, as such individuals may have attracted the attention of hunters and be more docile and easier to catch (Bried, 2003; Lees & Gilroy, 2009). Moreover, trading of penguins by Polynesians is unlikely, as only a few South Island endemic avian species have been found in fossil or archaeological deposits in the North Island (e.g. Worthy, 1998; Scofield *et al*., 2003), in comparison to cultural items (e.g. Davidson, 1988). Nevertheless, by the time Europeans arrived in NZ, *E. pachyrhynchus* was breeding only in the southwest of the South Island, an isolated region that has remained relatively inaccessible, with Polynesians largely abandoning the area following the extinction of terrestrial and coastal megafauna (Waters *et al*., 2017). These factors may help explain the apparently temporally stable genetic history of this species.

The findings of our study might be consistent with the suggestion of Mattern and Long (2017) and Mattern and Wilson (2018), who commented that *E. pachyrhynchus* populations may not be declining as rapidly as previously indicated. Despite our finding of little change in mtDNA diversity across several centuries of human occupation and predation, ongoing monitoring will ensure future management and appropriate conservation decisions are made for *E. pachyrhynchus*.

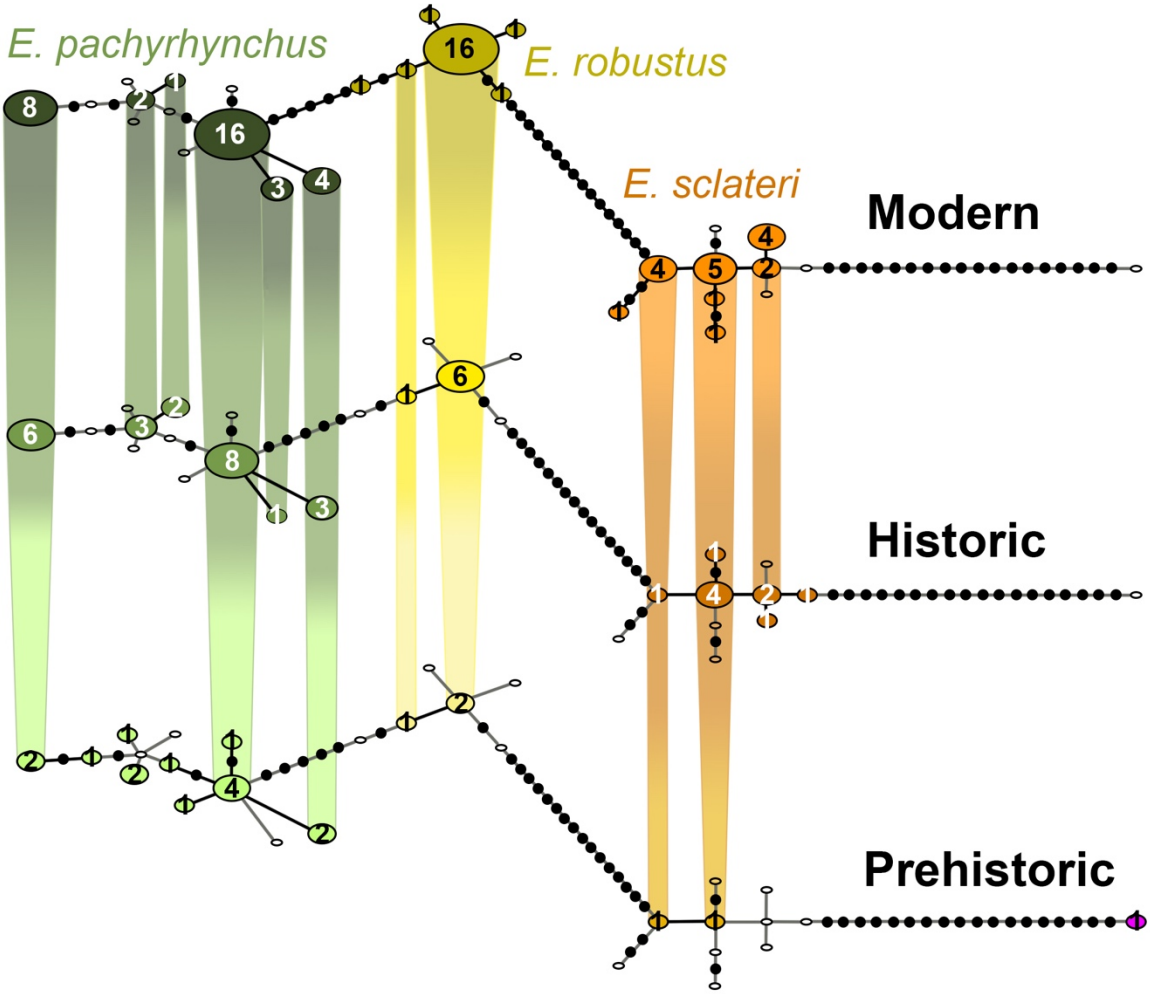
Conclusions

In conclusion, our aDNA analysis confirms the presence of substantial penguin species diversity in prehistoric mainland NZ, including the finding of *Eudyptes warhami cf*. *Eudyptes* clade X that has no extant equivalents. Despite a possible range contraction, the findings also suggest that the extant mainland *E. pachyrhynchus* has persisted within its core range, with relatively stable population size, likely owing to its presence in habitats that have experienced relatively limited human pressure. These findings highlight the potential for differential human impacts across species ranges.

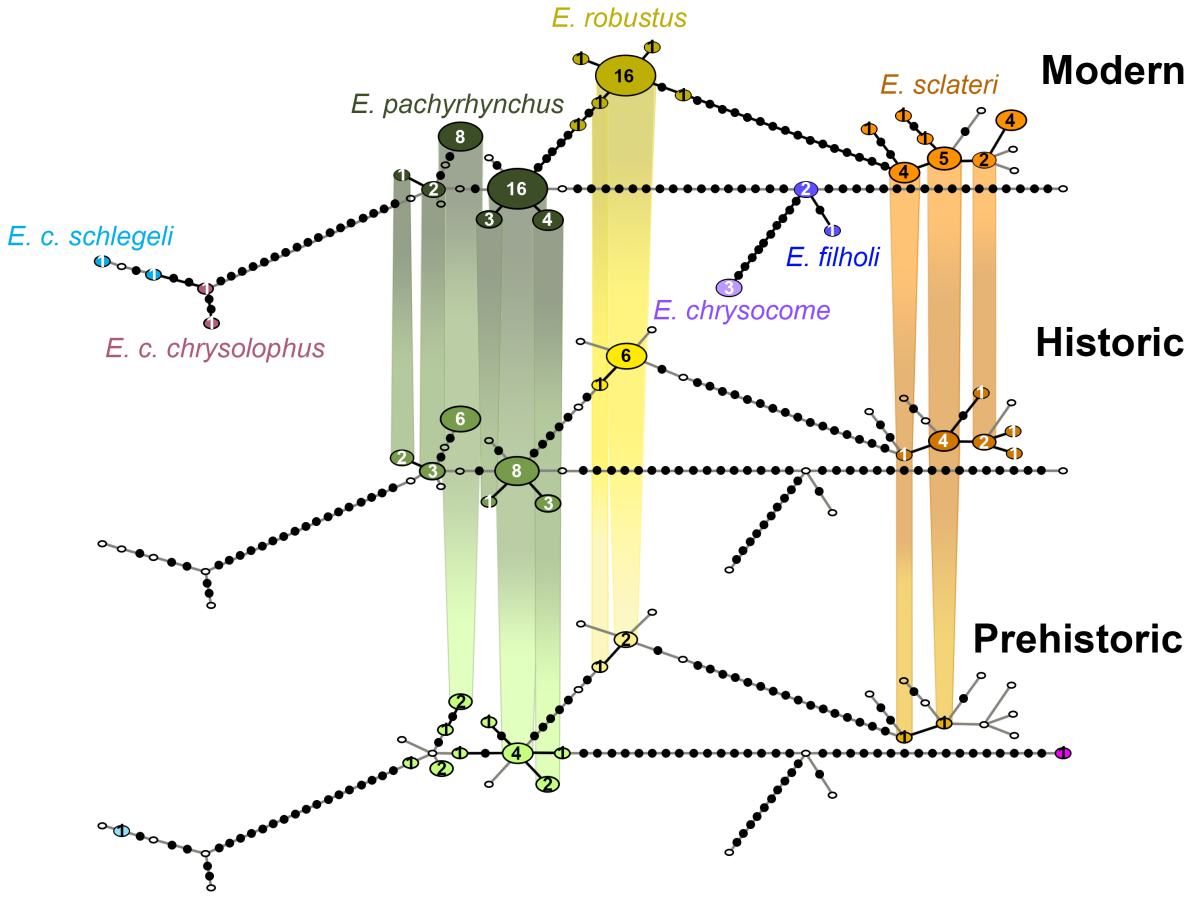
Supplementary Figures for *Chapter 4*



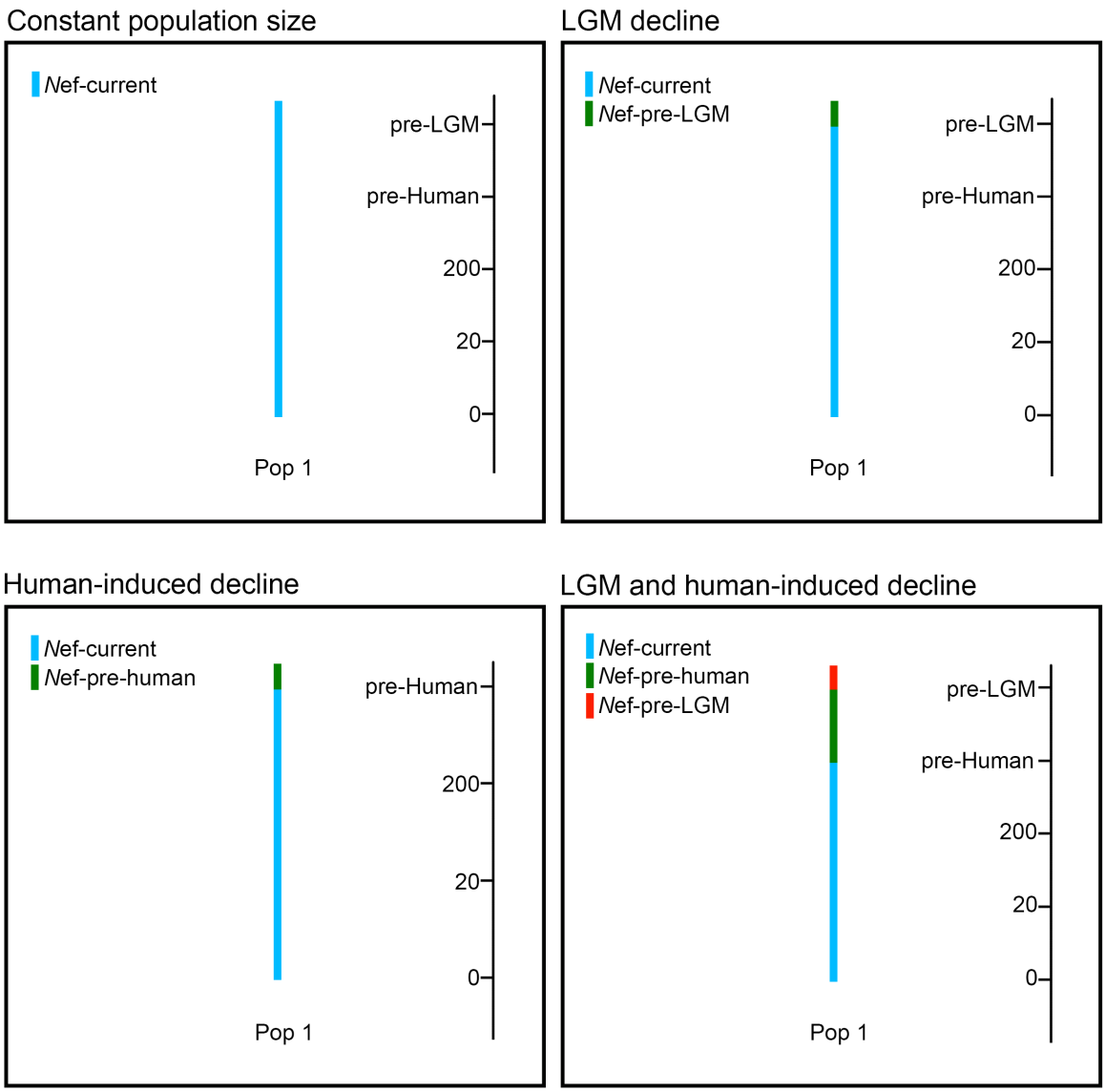
Supplementary Figure 22. Bayesian phylogenetic tree of penguins, with all aDNA COI sequences from NZ bone samples. *Eudyptes robustus* samples from the Western Chain are indicated with WC. AD numbers with the species name indicated were from historical museum skins, whereas samples without the species name (in bold) were were prehistoric subfossil or archaeological bone. Posterior probabilities for main clades >0.76 are shown.



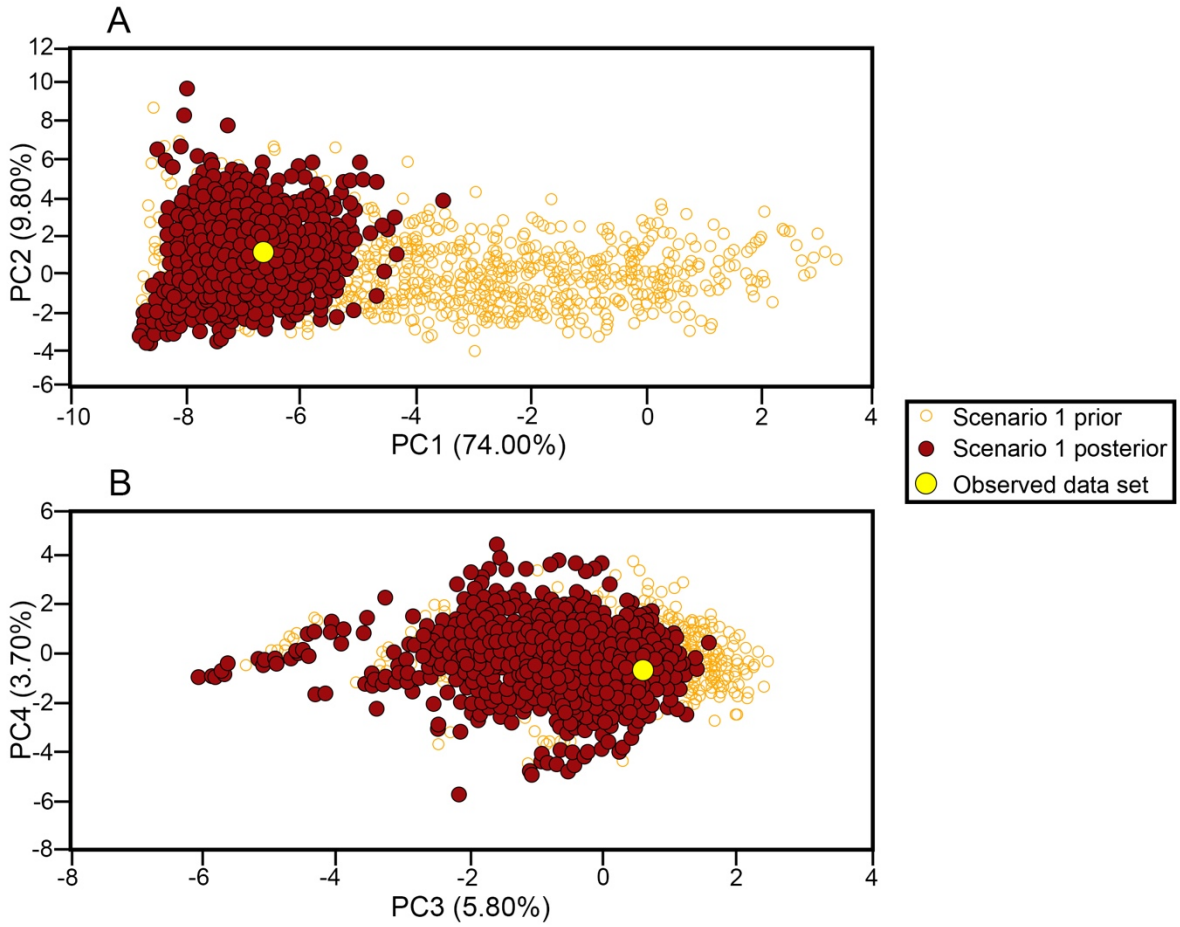
Supplementary Figure 23. Temporal haplotype network of *Eudyptes pachyrhynchus*, *E. robustus*, *E. sclateri* and *E. warhami*. Coloured circles represent different haplotypes, with the size of the circle and number within the circle corresponding to the number of samples that had that haplotype. Small white circles represent those haplotypes that are present in one or two-time periods, but are absent in the time period displayed. Black circles represent haplotypes that were absent from the dataset. Each layer represents a different time period, based on the sample type used (Prehistoric is sub-fossil and archaeological bone, Historic is historical museum skins and Modern is contemporary blood). The coloured cylinders connect those haplotypes that are present through multiple time periods. Green circles are *E. pachyrhynchus*, yellow circles are *E. robustus* and orange circles are *E. sclateri*. The single prehistoric pink circle is *Eudyptes warhami*/*Eudyptes* clade X.



Supplementary Figure 24. Temporal haplotype network of *Eudyptes* penguins. Coloured circles represent different haplotypes, with the size of the circle and number within the circle corresponding to the number of samples that had that haplotype. Small white circles represent those haplotypes that are present in one or two-time periods, but are absent in the time period displayed. Black circles represent haplotypes that were absent from the dataset. Each layer represents a different time period, based on the sample type used (Prehistoric is sub-fossil and archaeological bone, Historic is historical museum skins and Modern is contemporary blood). The coloured cylinders connect those haplotypes that are present through multiple time periods. Green circles are *E. pachyrhynchus*, yellow circles are *E. robustus*, orange circles are *E. sclateri*, light blue circles are *E. chrysolophus schlegeli*, mauve circles are *E. c. chrysolophus*, dark blue circles are *E. filholi*, the purple circle is *E. chrysocome* and the pink circle is *E. warhami* cf. *Eudyptes* clade X.



Supplementary Figure 25. Demographic models for *Eudyptes pachyrhynchus* associated with post-LGM climate change and human arrival in NZ. Time 0, 20 and 200 correspond to approximate sampling times.



Supplementary Figure 26. Model checking of the demographic analyses of *Eudyptes pachyrhynchus* associated with post-LGM climate change and human arrival in NZ. The figure shows a PCoA in the space of summary statistics using datasets simulated with the prior distributions of parameters (orange circles), observed dataset (yellow circle) and posterior distribution datasets (red circles). The model fits well with the observed data because the observed dataset (yellow circle) is in the middle of datasets from the posterior predictive distribution. **(A)** is PCs 1 and 2; and **(B)** is PCs 3 and 4.

Supplementary Table 35. Samples used for obtaining COI and CR. 'X' indicates that the marker of interest did not amplify, NA means the PCR was not attempted for that sample, GB indicates that COI sequences from that sample has been previously obtained (in Cole *et al*. 2018; *Chapter 5*), but are not analysed in this study. DNA was extracted twice from samples with two JW numbers.

Sample Number	Genus	Taxon	COI (bp)	Genbank Accession Number	*Eudyptes* CR (bp)	Genbank Accession Number	*Megadyptes* CR (bp)	Genbank Accession Number	Sample Type	Museum Number	Locality	Sample Type	Age (2016 is present)
JW718	*Eudyptes*	*chrysocome*	NA	NA	131	MK120620	NA	NA	Blood	370M	New Island, Falkland Islands	Wild	4 BCE
JW719	*Eudyptes*	*chrysocome*	NA	NA	131	MK120621	NA	NA	Blood	382F	New Island, Falkland Islands	Wild	4 BCE
JW720	*Eudyptes*	*chrysocome*	NA	NA	131	MK120622	NA	NA	Blood	374M	New Island, Falkland Islands	Wild	4 BCE
JW758	*Eudyptes*	*chrysolophus chrysolophus*	NA	NA	131	MK120623	NA	NA	Blood	MPJJ 1	Jardin Japonaise, Crozet	Wild	13 BCE
JW759	*Eudyptes*	*chrysolophus chrysolophus*	NA	NA	131	MK120624	NA	NA	Blood	MPJJ 2	Jardin Japonaise, Crozet	Wild	13 BCE
JW526	*Eudyptes*	*filholi*	NA	NA	131	MK120625	NA	NA	Blood	PIT 153746471 GLS #018 Male	Campbell Island	Wild	3 BCE
JW527	*Eudyptes*	*filholi*	NA	NA	131	MK120626	NA	NA	Blood	0153723532 GLS #010 Male	Campbell Island	Wild	3 BCE
JW528	*Eudyptes*	*filholi*	NA	NA	131	MK120627	NA	NA	Blood	167856729 GLS #8 Female	Campbell Island	Wild	3 BCE
AD107	*Eudyptes*	*pachyrhynchus*	GB	See *Supplementary Table 37*	131	MK120628	NA	NA	Bone	ANWC B21142	Hunter Island, Tasmania	Archaeological	461 cal years BP
AD108	*Eudyptes*	*pachyrhynchus*	GB	See *Supplementary Table 37*	131	MK120629	NA	NA	Bone	ANWC B21143	Hunter Island, Tasmania	Archaeological	461 cal years BP
AD155	*Eudyptes*	*pachyrhynchus*	GB	See *Supplementary Table 37*	131	MK120630	NA	NA	Bone	ANWC B21145	Louisa River Cave, Tasmania	Archaeological	541 cal years BP
AD257	*Eudyptes*	*pachyrhynchus*	GB	See *Supplementary Table 37*	131	MK120631	NA	NA	Toe-pad	OR.26082	Dusky Sound, Southland	Colony	32 BCE

Sample Number	Genus	Taxon	COI (bp)	Genbank Accession Number	*Eudyptes* CR (bp)	Genbank Accession Number	*Megadyptes* CR (bp)	Genbank Accession Number	Sample Type	Museum Number	Locality	Sample Type	Age (2016 is present)
AD259	*Eudyptes*	*pachyrhynchus*	GB	See *Supplementary Table 37*	131	MK120632	NA	NA	Toe-pad	OR.15371	Jackson Head, Westland	Colony	47 BCE
AD260	*Eudyptes*	*pachyrhynchus*	GB	*Supplementary Table 37*	131	MK120633	NA	NA	Toe-pad	OR.15368	Jackson Head, Westland	Colony	47 BCE
AD261	*Eudyptes*	*pachyrhynchus*	GB	See *Supplementary Table 37*	131	MK120634	NA	NA	Toe-pad	OR.23566	Turnbull River Mouth, Westland	Colony	31 BCE
AD262	*Eudyptes*	*pachyrhynchus*	GB	See *Supplementary Table 37*	131	MK120635	NA	NA	Toe-pad	OR.26575	Bathing Beach, Southland	Colony	28 BCE
AD264	*Eudyptes*	*pachyrhynchus*	GB	See *Supplementary Table 37*	131	MK120636	NA	NA	Toe-pad	OR.15835	Jackson Head, Westland	Colony	46 BCE
AD265	*Eudyptes*	*pachyrhynchus*	GB	See *Supplementary Table 37*	131	MK120637	NA	NA	Toe-pad	OR.15370	Jackson Head, Westland	Colony	47 BCE
AD266	*Eudyptes*	*pachyrhynchus*	GB	See *Supplementary Table 37*	131	MK120638	NA	NA	Toe-pad	OR.1067	Solander Island, Southland	Colony	68 BCE
AD267	*Eudyptes*	*pachyrhynchus*	GB	See *Supplementary Table 37*	131	MK120639	NA	NA	Toe-pad	OR.15369	Jackson Head, Westland	Colony	47 BCE
AD247	*Eudyptes*	*pachyrhynchus*	499	MK120774	131	MK120640	NA	NA	Bone	Av-36090	Cannibal Bay, Otago	Fossil/Archaeological	829 cal years BP
AD244	*Eudyptes*	*pachyrhynchus*	490	MK120775	131	MK120641	NA	NA	Bone	Av-22264	Preservation Inlet, Southland	Fossil	48 BCE
AD240	*Eudyptes*	*pachyrhynchus*	467	MK120776	131	MK120642	NA	NA	Bone	Av-34136	Pounawea, Otago	Archaeological	687 cal years BP
AD295	*Eudyptes*	*pachyrhynchus*	X	NA	131	MK120643	NA	NA	Bone	S.38968	Mussel Point, Marlborough	Archaeological	688 cal years BP
AD241	*Eudyptes*	*pachyrhynchus*	356	MK120778	131	MK120644	NA	NA	Bone	Av-34090	Pounawea, Otago	Archaeological	687 cal years BP
AD242	*Eudyptes*	*pachyrhynchus*	356	MK120779	131	MK120645	NA	NA	Bone	Av-34112	Pounawea, Otago	Archaeological	687 cal years BP
AD317	*Eudyptes*	*pachyrhynchus*	356	MK120780	131	MK120646	NA	NA	Bone	S.36369.7	Lake Grassmere, Marlborough	Fossil	1151 cal years BP
AD318	*Eudyptes*	*pachyrhynchus*	356	MK120781	131	MK120647	NA	NA	Bone	U13-L1	Foxton, Manawatu-Wanganui	Archaeological	651 cal years BP
AD357	*Eudyptes*	*pachyrhynchus*	283	MK120782	131	MK120648	NA	NA	Bone	S.23882	North Cape, Northland	Fossil	0-6000 BCE
AD340	*Eudyptes*	*pachyrhynchus*	263	MK120783	131	MK120649	NA	NA	Bone	Av-11170	Lake Grassmere, Marlborough	Fossil	1151 cal years BP

Sample Number	Genus	Taxon	COI (bp)	Genbank Accession Number	*Eudyptes* CR (bp)	Genbank Accession Number	*Megadyptes* CR (bp)	Genbank Accession Number	Sample Type	Museum Number	Locality	Sample Type	Age (2016 is present)
AD310	*Eudyptes*	*pachyrhynchus*	142	MK120784	X	NA	NA	NA	Bone	S.38443	Whakataki Beach, Wellington	Fossil	0 – 6000 BCE
AD311	*Eudyptes*	*pachyrhynchus*	142	MK120785	X	NA	NA	NA	Bone	S.38605	Ocean Beach, Hawkes Bay	Fossil	0 – 6000 BCE
AD360	*Eudyptes*	*pachyrhynchus*	142	MK120786	X	NA	NA	NA	Bone	S.45819	Te Kaukau Point, Wellington	Fossil	0-2000 BCE
JW130	*Eudyptes*	*pachyrhynchus*	458	MK120787	X	NA	NA	NA	Tissue	FCP1	Westport, Westland	Wild	3 BCE
AD263	*Eudyptes*	*pachyrhynchus*	426	MK120788	131	MK120650	NA	NA	Toe-pad	OR.15834	Jackson Head, Westland	Colony	46 BCE
AD298	*Eudyptes*	*pachyrhynchus*	356	MK120789	131	MK120651	NA	NA	Toe-pad	12501 (Auckland)	Puysegur Point, Southland	Colony	82 BCE
AD258	*Eudyptes*	*pachyrhynchus*	142	MK120790	131	MK120652	NA	NA	Toe-pad	OR.26083	Dusky Sound, Southland	Colony	32 BCE
AD287	*Eudyptes*	*pachyrhynchus*	X	NA	131	MK120653	NA	NA	Bone	S.36369.1	Lake Grassmere, Marlborough	Fossil	1151 cal years BP
AD243	*Eudyptes*	*pachyrhynchus*	X	NA	131	MK120654	NA	NA	Bone	Av-34866	False Islet, Otago	Fossil/Archaeological	829 cal years BP
AD268	*Eudyptes*	*pachyrhynchus*	NA	NA	131	MK120655	NA	NA	Toe-pad	OR.28064	Murphy Beach, Westland	Beach-wreck	16 BCE
AD320	*Eudyptes*	*pachyrhynchus*	NA	NA	131	MK120656	NA	NA	Toe-pad	Av-29716	South Beach, Greymouth, Westland	Beach-wreck	45 BCE
AD321	*Eudyptes*	*pachyrhynchus*	NA	NA	131	MK120657	NA	NA	Toe-pad	Av-17448	Punakaiki, Westland	Beach-wreck	18 BCE
AD322	*Eudyptes*	*pachyrhynchus*	NA	NA	131	MK120658	NA	NA	Toe-pad	Av-1363	Greymouth, Westland	Beach-wreck	77 BCE
AD325	*Eudyptes*	*pachyrhynchus*	NA	NA	131	MK120659	NA	NA	Toe-pad	Av-814	Hokitika, Westland	Beach-wreck	86 BCE
AD326	*Eudyptes*	*pachyrhynchus*	NA	NA	131	MK120660	NA	NA	Toe-pad	Av-44	Solander Island, Southland	Colony	85 BCE
AD334	*Eudyptes*	*pachyrhynchus*	NA	NA	131	MK120661	NA	NA	Toe-pad	Av-36574	Okarito, Westland	Beach-wreck	39 BCE
AD335	*Eudyptes*	*pachyrhynchus*	NA	NA	131	MK120662	NA	NA	Toe-pad	Av-22735	Rapahoe Beach, Westland	Beach-wreck	35 BCE
JW348	*Eudyptes*	*pachyrhynchus*	NA	NA	131	MK120663	NA	NA	Tissue	FCP2	Westport, Westland	Wild	3 BCE
JW349	*Eudyptes*	*pachyrhynchus*	NA	NA	131	MK120664	NA	NA	Tissue	FCP3	Westport, Westland	Wild	3 BCE

Sample Number	Genus	Taxon	COI (bp)	Genbank Accession Number	*Eudyptes* CR (bp)	Genbank Accession Number	*Megadyptes* CR (bp)	Genbank Accession Number	Sample Type	Museum Number	Locality	Sample Type	Age (2016 is present)
JW350	*Eudyptes*	*pachyrhynchus*	NA	NA	131	MK120665	NA	NA	Tissue	FCP4	Murphy's Beach, Westland	Wild	3 BCE
JW351	*Eudyptes*	*pachyrhynchus*	NA	NA	131	MK120666	NA	NA	Tissue	FCP5	Murphy's Beach, Westland	Wild	3 BCE
JW412	*Eudyptes*	*pachyrhynchus*	NA	NA	131	MK120667	NA	NA	Tissue	LB-12810	Brighton Beach, Otago	Wild	11 BCE
JW414	*Eudyptes*	*pachyrhynchus*	NA	NA	131	MK120668	NA	NA	Blood	LB-12809	Okarito, Westland	Wild	12 BCE
AD323	*Eudyptes*	*pachyrhynchus*	NA	NA	131	MK120669	NA	NA	Skin	Av-850	Preservation Inlet, Southland	Colony	95 BCE
AD324	*Eudyptes*	*pachyrhynchus*	NA	NA	131	MK120670	NA	NA	Feather	Av-417	Coal Island, Preservation Inlet, Southland	Colony	96 BCE
JW319	*Eudyptes*	*pachyrhynchus*	NA	NA	131	MK120671	NA	NA	Blood	WH31	Codfish Island, Southland	Wild	0 BCE
JW334	*Eudyptes*	*pachyrhynchus*	NA	NA	131	MK120672	NA	NA	Blood	JH2016	Jackson Head, Westland	Wild	0 BCE
JW337	*Eudyptes*	*pachyrhynchus*	NA	NA	131	MK120673	NA	NA	Blood	MS2	Milford Sound, Southland	Wild	0 BCE
JW338	*Eudyptes*	*pachyrhynchus*	NA	NA	131	MK120674	NA	NA	Blood	MS3	Milford Sound, Southland	Wild	0 BCE
JW343	*Eudyptes*	*pachyrhynchus*	NA	NA	131	MK120675	NA	NA	Blood	MS8	Milford Sound, Southland	Wild	0 BCE
JW344	*Eudyptes*	*pachyrhynchus*	NA	NA	131	MK120676	NA	NA	Blood	MS9	Milford Sound, Southland	Wild	0 BCE
JW345	*Eudyptes*	*pachyrhynchus*	NA	NA	131	MK120677	NA	NA	Blood	MS10	Milford Sound, Southland	Wild	0 BCE
JW420	*Eudyptes*	*pachyrhynchus*	NA	NA	131	MK120678	NA	NA	Blood	WH01	Codfish Island, Southland	Wild	0 BCE
JW421	*Eudyptes*	*pachyrhynchus*	NA	NA	131	MK120679	NA	NA	Blood	WH02	Codfish Island, Southland	Wild	0 BCE
JW422	*Eudyptes*	*pachyrhynchus*	NA	NA	131	MK120680	NA	NA	Blood	WH11	Codfish Island, Southland	Wild	0 BCE

Sample Number	Genus	Taxon	COI (bp)	Genbank Accession Number	*Eudyptes* CR (bp)	Genbank Accession Number	*Megadyptes* CR (bp)	Genbank Accession Number	Sample Type	Museum Number	Locality	Sample Type	Age (2016 is present)
JW424	*Eudyptes*	*pachyrhynchus*	NA	NA	131	MK120681	NA	NA	Blood	WH18	Codfish Island, Southland	Wild	0 BCE
JW432	*Eudyptes*	*pachyrhynchus*	NA	NA	131	MK120682	NA	NA	Blood	JH10	Jackson Head, Westland	Wild	0 BCE
JW433	*Eudyptes*	*pachyrhynchus*	NA	NA	131	MK120683	NA	NA	Blood	JH21	Jackson Head, Westland	Wild	0 BCE
JW441	*Eudyptes*	*pachyrhynchus*	NA	NA	131	MK120684	NA	NA	Blood	MS10	Milford Sound, Southland	Wild	0 BCE
JW639	*Eudyptes*	*pachyrhynchus*	NA	NA	131	MK120685	NA	NA	Blood	WH11	Codfish Island, Southland	Wild	0 BCE
JW645	*Eudyptes*	*pachyrhynchus*	NA	NA	131	MK120686	NA	NA	Blood	WH11A	Codfish Island, Southland	Wild	0 BCE
JW646	*Eudyptes*	*pachyrhynchus*	NA	NA	131	MK120687	NA	NA	Blood	WH13	Codfish Island, Southland	Wild	0 BCE
JW647	*Eudyptes*	*pachyrhynchus*	NA	NA	131	MK120688	NA	NA	Blood	WH16	Codfish Island, Southland	Wild	0 BCE
JW648	*Eudyptes*	*pachyrhynchus*	NA	NA	131	MK120689	NA	NA	Blood	WH17	Codfish Island, Southland	Wild	0 BCE
JW649	*Eudyptes*	*pachyrhynchus*	NA	NA	131	MK120690	NA	NA	Blood	WH18	Codfish Island, Southland	Wild	0 BCE
JW650	*Eudyptes*	*pachyrhynchus*	NA	NA	131	MK120691	NA	NA	Blood	WH28	Codfish Island, Southland	Wild	0 BCE
JW651	*Eudyptes*	*pachyrhynchus*	NA	NA	131	MK120692	NA	NA	Blood	JH01	Jackson Head, Westland	Wild	0 BCE
JW656	*Eudyptes*	*pachyrhynchus*	NA	NA	131	MK120693	NA	NA	Blood	JH21	Jackson Head, Westland	Wild	0 BCE
JW658	*Eudyptes*	*pachyrhynchus*	NA	NA	131	MK120694	NA	NA	Blood	JH Cave 2016	Jackson Head, Westland	Wild	0 BCE
JW659	*Eudyptes*	*pachyrhynchus*	NA	NA	131	MK120695	NA	NA	Blood	JH 2016	Jackson Head, Westland	Wild	0 BCE
JW665	*Eudyptes*	*pachyrhynchus*	NA	NA	131	MK120696	NA	NA	Blood	MS6	Milford Sound, Southland	Wild	0 BCE

Sample Number	Genus	Taxon	COI (bp)	Genbank Accession Number	*Eudyptes* CR (bp)	Genbank Accession Number	*Megadyptes* CR (bp)	Genbank Accession Number	Sample Type	Museum Number	Locality	Sample Type	Age (2016 is present)
JW667	*Eudyptes*	*pachyrhynchus*	NA	NA	131	MK120697	NA	NA	Blood	MS10	Milford Sound, Southland	Wild	0 BCE
JW668	*Eudyptes*	*pachyrhynchus*	NA	NA	131	MK120698	NA	NA	Blood	MS11	Milford Sound, Southland	Wild	0 BCE
AD154	*Eudyptes*	*robustus*	GB	See *Supplementary Table 37*	131	MK120699	NA	NA	Bone	ANWC B21144	Prion Beach, Tasmania	Archaeological	Unknown
AD245	*Eudyptes*	*robustus*	499	MK120793	131	MK120700	NA	NA	Bone	Av-28286	Native Island, Southland	Fossil	1505 cal years BP
AD294	*Eudyptes*	*robustus*	356	MK120794	131	MK120701	NA	NA	Bone	S.36176	Mason Bay, Southland	Fossil	0 – 407 BCE
AD270	*Eudyptes*	*robustus*	GB	See *Supplementary Table 37*	131	MK120702	NA	NA	Toe-pad	OR.19353	The Snares	Colony	39 BCE
AD272	*Eudyptes*	*robustus*	GB	See *Supplementary Table 37*	131	MK120703	NA	NA	Toe-pad	OR.26830	Penguin Creek, The Snares	Colony	16 BCE
AD275	*Eudyptes*	*robustus*	169	MK120795	131	MK120704	NA	NA	Toe-pad	OR.17375	Western Chain, The Snares	Colony	44 BCE
AD273	*Eudyptes*	*robustus*	168	MK120796	131	MK120705	NA	NA	Toe-pad	OR.26831	Muttonbird Creek, The Snares	Colony	16 BCE
AD274	*Eudyptes*	*robustus*	171	MK120797	131	MK120706	NA	NA	Toe-pad	OR.26832	Sinkhole Flat, The Snares	Colony	16 BCE
AD271	*Eudyptes*	*robustus*	NA	NA	131	MK120707	NA	NA	Toe-pad	OR.24113	The Snares	Colony	29 BCE
AD319	*Eudyptes*	*robustus*	NA	NA	131	MK120708	NA	NA	Toe-pad	Av-804	Otago Coast, Otago	Beach-wreck	124 BCE
JW127/JW485	*Eudyptes*	*robustus*	462	MK120798	131	MK120709	NA	NA	Blood	SCP8	Toru, Western Chain	Wild	3 BCE
JW120/JW480	*Eudyptes*	*robustus*	459	MK120799	131	MK120710	NA	NA	Blood	SCP1	South Promontory	Wild	3 BCE
JW121/JW481	*Eudyptes*	*robustus*	459	MK120800	131	MK120711	NA	NA	Blood	SCP2	South Promontory	Wild	3 BCE
JW122/JW482	*Eudyptes*	*robustus*	459	MK120801	131	MK120712	NA	NA	Blood	SCP3	North Promontory	Wild	3 BCE
JW125/JW713	*Eudyptes*	*robustus*	459	MK120802	131	MK120713	NA	NA	Blood	SCP6	Colony 3	Wild	3 BCE
JW126/JW484	*Eudyptes*	*robustus*	459	MK120803	131	MK120714	NA	NA	Blood	SCP7	Toru, Western Chain	Wild	3 BCE
JW129	*Eudyptes*	*robustus*	459	MK120804	131	MK120715	NA	NA	Blood	SCP10	Toru, Western Chain	Wild	3 BCE

Sample Number	Genus	Taxon	COI (bp)	Genbank Accession Number	*Eudyptes* CR (bp)	Genbank Accession Number	*Megadyptes* CR (bp)	Genbank Accession Number	Sample Type	Museum Number	Locality	Sample Type	Age (2016 is present)
JW128	*Eudyptes*	*robustus*	458	MK120805	131	MK120716	NA	NA	Blood	SCP9	Toru, Western Chain	Wild	3 BCE
JW124/JW712	*Eudyptes*	*robustus*	457	MK120806	131	MK120717	NA	NA	Blood	SCP5	Colony 3	Wild	3 BCE
JW180	*Eudyptes*	*robustus*	NA	NA	131	MK120718	NA	NA	Blood	68M	The Snares	Wild	3 BCE
JW181	*Eudyptes*	*robustus*	NA	NA	131	MK120719	NA	NA	Blood	38F	The Snares	Wild	3 BCE
JW444	*Eudyptes*	*robustus*	NA	NA	131	MK120720	NA	NA	Blood	13F	The Snares	Wild	3 BCE
JW445	*Eudyptes*	*robustus*	NA	NA	131	MK120721	NA	NA	Blood	14F	The Snares	Wild	3 BCE
JW446	*Eudyptes*	*robustus*	NA	NA	131	MK120722	NA	NA	Blood	49M	The Snares	Wild	3 BCE
JW447	*Eudyptes*	*robustus*	NA	NA	131	MK120723	NA	NA	Blood	18M	The Snares	Wild	3 BCE
JW448	*Eudyptes*	*robustus*	NA	NA	131	MK120724	NA	NA	Blood	36M	The Snares	Wild	3 BCE
JW449	*Eudyptes*	*robustus*	NA	NA	131	MK120725	NA	NA	Blood	58M	The Snares	Wild	3 BCE
JW450	*Eudyptes*	*robustus*	NA	NA	131	MK120726	NA	NA	Blood	59F	The Snares	Wild	3 BCE
JW451	*Eudyptes*	*robustus*	NA	NA	131	MK120727	NA	NA	Blood	29M	The Snares	Wild	3 BCE
JW456	*Eudyptes*	*robustus*	NA	NA	131	MK120728	NA	NA	Blood	46 gvn F	The Snares	Wild	3 BCE
JW457	*Eudyptes*	*robustus*	NA	NA	131	MK120729	NA	NA	Blood	65F	The Snares	Wild	3 BCE
JW711	*Eudyptes*	*robustus*	NA	NA	131	MK120730	NA	NA	Blood	SCP4	North Promontory	Wild	3 BCE
AD367	*Eudyptes*	*schlegeli* or *chrysolophus*	356	MK120807	131	MK120731	NA	NA	Bone	S.37221	Native Island, Southland	Fossil	1505 cal years BP
JW869	*Eudyptes*	*schlegeli*	NA	NA	131	MK120732	NA	NA	Blood	ROPE4461	Green Gorge, Macquarie Island	Wild	10 BCE
JW870	*Eudyptes*	*schlegeli*	NA	NA	131	MK120733	NA	NA	Blood	ROPE4449	Green Gorge, Macquarie Island	Wild	10 BCE
AD344	*Eudyptes*	*sclateri*	352	MK120808	131	MK120734	NA	NA	Bone	S.47280	Sharks Tooth, Kaikoura, Canterbury	Fossil	0 – 6000 BCE
AD389	*Eudyptes*	*sclateri*	280	MK120809	131	MK120735	NA	NA	Bone	4/1/5 Black Rocks Cresent BR4 Box 2a 1 (Bird)	Black Rocks, Wairarapa, Wellington	Archaeological	733 – 566 Cal years BP

Sample Number	Genus	Taxon	COI (bp)	Genbank Accession Number	*Eudyptes* CR (bp)	Genbank Accession Number	*Megadyptes* CR (bp)	Genbank Accession Number	Sample Type	Museum Number	Locality	Sample Type	Age (2016 is present)
AD393	*Eudyptes*	*sclateri*	283	MK120810	X	NA	NA	NA	Bone	4/2/6 Black Rocks Cresent BR4 Box 2a 1 (Bird)	Black Rocks, Wairarapa, Wellington	Archaeological	733 – 566 Cal years BP
AD304	*Eudyptes*	*sclateri*	GB	See *Supplementary Table 37*	131	MK120736	NA	NA	Toe-pad	OR.21942	Antipodes Islands	Colony	48 BCE
AD302	*Eudyptes*	*sclateri*	499	MK120811	131	MK120737	NA	NA	Toe-pad	OR.12646	Campbell Island	Colony	69 BCE
AD303	*Eudyptes*	*sclateri*	NA	NA	131	MK120738	NA	NA	Toe-pad	OR.12647	Campbell Island	Colony	69 BCE
AD305	*Eudyptes*	*sclateri*	NA	NA	131	MK120739	NA	NA	Toe-pad	OR.21941	Antipodes Islands	Colony	48 BCE
AD327	*Eudyptes*	*sclateri*	NA	NA	131	MK120740	NA	NA	Toe-pad	Av-1353	Antipodes Islands	Colony	93 BCE
AD329	*Eudyptes*	*sclateri*	NA	NA	131	MK120741	NA	NA	Toe-pad	Av-808	Antipodes Islands	Colony	93 BCE
AD332	*Eudyptes*	*sclateri*	NA	NA	131	MK120742	NA	NA	Toe-pad	Av-809	Antipodes Islands	Colony	93 BCE
AD333	*Eudyptes*	*sclateri*	NA	NA	131	MK120743	NA	NA	Toe-pad	Av-21754	Campbell Island	Colony	67 BCE
AD306	*Eudyptes*	*sclateri*	NA	NA	131	MK120744	NA	NA	Skin	OR.26573	Anchorage Bay, Antipodes Islands	Colony	21 BCE
AD307	*Eudyptes*	*sclateri*	NA	NA	131	MK120745	NA	NA	Skin	OR.21944	Antipodes Islands	Colony	38 BCE
JW183	*Eudyptes*	*sclateri*	NA	NA	131	MK120746	NA	NA	Blood	Ant 48	Antipodes Islands	Wild	11 BCE
JW184	*Eudyptes*	*sclateri*	NA	NA	131	MK120747	NA	NA	Blood	Ant 5	Antipodes Islands	Wild	11 BCE
JW496	*Eudyptes*	*sclateri*	NA	NA	131	MK120748	NA	NA	Blood	Ant 6	Antipodes Islands	Wild	11 BCE
JW497	*Eudyptes*	*sclateri*	NA	NA	131	MK120749	NA	NA	Blood	Ant 16	Antipodes Islands	Wild	11 BCE
JW498	*Eudyptes*	*sclateri*	NA	NA	131	MK120750	NA	NA	Blood	Ant 4	Antipodes Islands	Wild	11 BCE
JW499	*Eudyptes*	*sclateri*	NA	NA	131	MK120751	NA	NA	Blood	Ant 7	Antipodes Islands	Wild	11 BCE
JW500	*Eudyptes*	*sclateri*	NA	NA	131	MK120752	NA	NA	Blood	Ant 3	Antipodes Islands	Wild	11 BCE
JW501	*Eudyptes*	*sclateri*	NA	NA	131	MK120753	NA	NA	Blood	Ant 46	Antipodes Islands	Wild	11 BCE
JW502	*Eudyptes*	*sclateri*	NA	NA	131	MK120754	NA	NA	Blood	Ant 45	Antipodes Islands	Wild	11 BCE
JW503	*Eudyptes*	*sclateri*	NA	NA	131	MK120755	NA	NA	Blood	Ant 43	Antipodes Islands	Wild	11 BCE
JW504	*Eudyptes*	*sclateri*	NA	NA	131	MK120756	NA	NA	Blood	Ant 42	Antipodes Islands	Wild	11 BCE

Sample Number	Genus	Taxon	COI (bp)	Genbank Accession Number	*Eudyptes* CR (bp)	Genbank Accession Number	*Megadyptes* CR (bp)	Genbank Accession Number	Sample Type	Museum Number	Locality	Sample Type	Age (2016 is present)
JW505	*Eudyptes*	*sclateri*	NA	NA	131	MK120757	NA	NA	Blood	Ant 15	Antipodes Islands	Wild	11 BCE
JW506	*Eudyptes*	*sclateri*	NA	NA	131	MK120758	NA	NA	Blood	Ant 13	Antipodes Islands	Wild	11 BCE
JW507	*Eudyptes*	*sclateri*	NA	NA	131	MK120759	NA	NA	Blood	Ant 44	Antipodes Islands	Wild	11 BCE
JW508	*Eudyptes*	*sclateri*	NA	NA	131	MK120760	NA	NA	Blood	Ant 8	Antipodes Islands	Wild	11 BCE
JW509	*Eudyptes*	*sclateri*	NA	NA	131	MK120761	NA	NA	Blood	Ant 14	Antipodes Islands	Wild	11 BCE
JW510	*Eudyptes*	*sclateri*	NA	NA	131	MK120762	NA	NA	Blood	Ant 10	Antipodes Islands	Wild	11 BCE
JW511	*Eudyptes*	*sclateri*	NA	NA	131	MK120763	NA	NA	Blood	Ant 12	Antipodes Islands	Wild	11 BCE
AD391	*Eudyptes*	*warhami* cf. clade X	499	MK120812	111	MK120764	NA	NA	Bone	4/2/3 Black Rocks Cresent BR4 Box 2a 1 (Bird)	Black Rocks, Wairarapa, Wellington	Archaeological	733 – 566 Cal years BP
AD342	*Eudyptes*	*warhami* cf. clade X	499	MK120813	X	NA	NA	NA	Bone	Av-25895	Le Bons Bay, Canterbury	Fossil	0 – 6000 BCE
AD309	*Eudyptes*	*warhami* cf. clade X	356	MK120814	X	NA	NA	NA	Bone	3/2/8 Black Rocks Midden BR3 N168-9/77	Black Rocks, Wairarapa, Wellington	Archaeological	733 – 566 Cal years BP
AD248	*Eudyptes*	*warhami* cf. clade X	286	MK120815	X	NA	NA	NA	Bone	Av-21751	Rakauteru Cave, Kaikoura Canterbury	Archaeological	733 – 566 Cal years BP
AD381	*Eudyptes*	*warhami* cf. clade X	282	MK120816	X	NA	NA	NA	Bone	Av.32062	Needle Point North, Marlborough	Archaeological	733 – 566 Cal years BP
AD394	*Eudyptes*	*warhami* cf. clade X	274	MK120817	X	NA	NA	NA	Bone	4/1/6 Black Rocks Cresent BR4 Box 2a 1 (Bird)	Black Rocks, Wairarapa, Wellington	Archaeological	733 – 566 Cal years BP
AD282	*Eudyptes*	*warhami* cf. clade X	251	MK120818	X	NA	NA	NA	Bone	M1/II/4-60 (rust + 5r)	Centennial Inn, Paekakariki, Wellington	Archaeological	733 – 566 Cal years BP
AD93	*Megadyptes*	*antipodes* ssp.	377	MK120819	NA	NA	GB	See *Supplementary Table 37*	Bone	Otago UF2/L2, LB3 A6 bird L2+Llb Bag LB/D/F	Long Beach, Otago	Archaeological	541 – 351 cal years BP
AD96	*Megadyptes*	*antipodes* ssp.	204	MK120820	NA	NA	GB	See *Supplementary Table 37*	Bone	Av-10616	Pahia Pa, Southland	Archaeological	471 – 192 cal years BP
AD94	*Megadyptes*	*antipodes* ssp.	141	MK120821	NA	NA	GB	See *Supplementary Table 37*	Bone	Otago UE3/L2	Long Beach, Otago	Archaeological	441 – 151 cal years BP

Sample Number	Genus	Taxon	COI (bp)	Genbank Accession Number	*Eudyptes* CR (bp)	Genbank Accession Number	*Megadyptes* CR (bp)	Genbank Accession Number	Sample Type	Museum Number	Locality	Sample Type	Age (2016 is present)
AD91	*Megadyptes*	*antipodes waitaha*	141	MK120822	NA	NA	GB	See *Supplementary Table 37*	Bone	S.35913	Native Island, Southland	Fossil	1507 cal years BP
AD92	*Megadyptes*	*antipodes waitaha*	138	MK120823	NA	NA	GB	See *Supplementary Table 37*	Bone	S.35426	Native Island, Southland	Fossil	1507 cal years BP
AD284	*Megadyptes*	*antipodes waitaha*	213	MK120824	NA	NA	402	MK120765	Bone	Moa hunter midden, Old area, October 62, Left humerus small	Sinclair, Paremata, Wellington	Archaeological	733 – 566 Cal years BP
AD286	*Megadyptes*	*antipodes waitaha*	141	MK120825	NA	NA	402	MK120766	Bone	Yellow Surface	Mana Island North, Wellington	Archaeological	733 – 566 Cal years BP
AD378	*Megadyptes*	*antipodes waitaha*	141	MK120826	NA	NA	402	MK120767	Bone	S.38481	Delaware Bay, Nelson	Fossil	1330 – 918 cal years BP
AD283	*Megadyptes*	*antipodes waitaha*	137	MK120827	NA	NA	402	MK120768	Bone	Moa hunter midden, Old area, October 62, Left humerus large	Sinclair, Paremata, Wellington	Archaeological	733 – 566 Cal years BP
AD289	*Megadyptes*	*antipodes waitaha*	109	MK120828	NA	NA	402	MK120769	Bone	S.42136	Delaware Bay, Nelson	Fossil	1330 – 918 cal years BP
AD285	*Megadyptes*	*antipodes waitaha*	233	MK120829	NA	NA	255	MK120770	Bone	N-160/50	Sinclair, Paremata, Wellington	Archaeological	733 – 566 Cal years BP
AD371	*Megadyptes*	*antipodes waitaha*	141	MK120830	NA	NA	255	MK120771	Bone	Av-34054	Pounawea, Otago	Archaeological	687 cal years BP
AD375	*Megadyptes*	*antipodes waitaha*	X	NA	NA	NA	242	MK120772	Bone	Av-34866	False Islet, Otago	Fossil/Archaeological	829 cal years BP
AD276	*Megadyptes*	*antipodes waitaha*	252	MK120831	NA	NA	402	MK120773	Bone	Ar-83	Centennial Inn, Paekakariki, Wellington	Archaeological	733 – 566 Cal years BP
AD379	*Megadyptes*	*antipodes* ssp.	139	MK120832	NA	NA	X	NA	Bone	S.42118	Whangamoa River Mouth, Nelson	Fossil	0 – 6000 BP
AD366	*Megadyptes*	*antipodes* ssp.	141	MK120833	NA	NA	X	NA	Bone	Av-36083	False Islet, Otago	Fossil/Archaeological	829 cal years BP
AD358	*Megadyptes*	*antipodes* ssp.	130	MK120834	NA	NA	X	NA	Bone	S.41881.1	Mussel Point, Marlborough	Archaeological	733 – 566 Cal years BP
AD249	X	X	X	NA	X	NA	X	NA	Bone	Av27268	NE Bay, Solander Island, Southland	Fossil	NA

Sample Number	Genus	Taxon	COI (bp)	Genbank Accession Number	*Eudyptes* CR (bp)	Genbank Accession Number	*Megadyptes* CR (bp)	Genbank Accession Number	Sample Type	Museum Number	Locality	Sample Type	Age (2016 is present)
AD279	X	X	X	NA	X	NA	X	NA	Bone	M1/11/J/L4 Black	Centennial Inn, Paekakariki, Wellington	Archaeological	733 – 566 Cal years BP
AD280	X	X	X	NA	X	NA	X	NA	Bone	M1/2/23/Lens 2B	Centennial Inn, Paekakariki, Wellington	Archaeological	733 – 566 Cal years BP
AD281	X	X	X	NA	X	NA	X	NA	Bone	M1/2/9 Lens B	Centennial Inn, Paekakariki, Wellington	Archaeological	733 – 566 Cal years BP
AD288	X	X	X	NA	X	NA	X	NA	Bone	S.36369.2	Lake Grassmere, Marlborough	Fossil	1151 cal years BP
AD312	X	X	X	NA	X	NA	X	NA	Bone	S.45814	White Rock, Wellington	Fossil	0 – 2000 BCE
AD313	X	X	X	NA	X	NA	X	NA	Bone	S.45821	Te Kaukau Point, Wellington	Fossil	0 – 2000 BCE
AD314	X	X	X	NA	X	NA	X	NA	Bone	S.36369.3	Lake Grassmere, Marlborough	Fossil	1151 cal years BP
AD315	X	X	X	NA	X	NA	X	NA	Bone	S.36369.4	Lake Grassmere, Marlborough	Fossil	1151 cal years BP
AD316	X	X	X	NA	X	NA	X	NA	Bone	S.36369.5	Lake Grassmere, Marlborough	Fossil	1151 cal years BP
AD337	X	X	X	NA	X	NA	X	NA	Bone	Av-14445	Cave, Punakaiki, Westland	Fossil	93 cal years BP
AD338	X	X	X	NA	X	NA	X	NA	Bone	Av-14444	Limestone Cave, Punakaiki, Westland	Fossil	n/a
AD339	X	X	X	NA	X	NA	X	NA	Bone	Av-11568	Marfells Beach, Marlborough	Fossil	1151 cal years BP
AD353	X	X	X	NA	X	NA	X	NA	Bone	Av-34090	Pounawea, Otago	Archaeological	687 cal years BP
AD355	X	X	X	NA	X	NA	X	NA	Bone	Av-34112	Pounawea, Otago	Archaeological	687 cal years BP
AD356	X	X	X	NA	X	NA	X	NA	Bone	4/1/6 Black Rocks Crescent BR4 Box 2a1 (bird)	Black Rocks, Wairarapa, Wellington	Archaeological	733 – 566 Cal years BP

Sample Number	Genus	Taxon	COI (bp)	Genbank Accession Number	*Eudyptes* CR (bp)	Genbank Accession Number	*Megadyptes* CR (bp)	Genbank Accession Number	Sample Type	Museum Number	Locality	Sample Type	Age (2016 is present)
AD359	X	X	X	NA	X	NA	X	NA	Bone	S.45820	Te Kaukau Point, Wellington	Fossil	0 – 2000 BCE
AD361	X	X	X	NA	X	NA	X	NA	Bone	S.24724	Paremata, Wellington	Archaeological	733 – 566 Cal years BP
AD368	X	X	X	NA	X	NA	X	NA	Bone	S.35912	Native Island, Southland	Fossil	1505 cal years BP
AD369	X	X	X	NA	X	NA	X	NA	Bone	Av-21257	Cape Wanbrow, Otago	Fossil	>10,000 BCE
AD372	X	X	X	NA	X	NA	X	NA	Bone	Av-34112	Pounawea, Otago	Archaeological	687 cal years BP
AD373	X	X	X	NA	X	NA	X	NA	Bone	Av-34054	Pounawea, Otago	Archaeological	687 cal years BP
AD374	X	X	X	NA	X	NA	X	NA	Bone	Av-32062	North Needle Point, Marlborough	Archaeological	733 – 566 Cal years BP
AD380	X	X	X	NA	X	NA	X	NA	Bone	Av-34054	Pounawea, Otago	Archaeological	687 cal years BP
AD382	X	X	X	NA	X	NA	X	NA	Bone	S.41881.2	Mussel Point, Marlborough	Archaeological	733 – 566 Cal years BP
AD383	X	X	X	NA	X	NA	X	NA	Bone	Av-14214	Marfells Beach, Marlborough	Fossil	1151 cal years BP
AD384	X	X	X	NA	X	NA	X	NA	Bone	S.38549	Delaware Bay, Nelson	Fossil	1330 – 918 cal years BP
AD386	X	X	X	NA	X	NA	X	NA	Bone	Av-190	Paremata, Wellington	Archaeological	733 – 566 Cal years BP
AD387	X	X	X	NA	X	NA	X	NA	Bone	4/3/8 Black Rocks Cresent BR4 Box 2a 1 (Bird)	Black Rocks, Wairarapa, Wellington	Archaeological	733 – 566 Cal years BP
AD388	X	X	X	NA	X	NA	X	NA	Bone	4/3/1 Black Rocks Cresent BR4 Box 2a 1 (Bird)	Black Rocks, Wairarapa, Wellington	Archaeological	733 – 566 Cal years BP
AD390	X	X	X	NA	X	NA	X	NA	Bone	3/1/4 Black Rocks Black Midden BR3 N168-9/77	Black Rocks, Wairarapa, Wellington	Archaeological	733 – 566 Cal years BP

Sample Number	Genus	Taxon	COI (bp)	Genbank Accession Number	*Eudyptes* CR (bp)	Genbank Accession Number	*Megadyptes* CR (bp)	Genbank Accession Number	Sample Type	Museum Number	Locality	Sample Type	Age (2016 is present)
AD392	X	X	X	NA	X	NA	X	NA	Bone	4/2/4 Black Rocks Cresent BR4 Box 2a 1 (Bird)	Black Rocks, Wairarapa, Wellington	Archaeological	733 – 566 Cal years BP
AD395	X	X	X	NA	X	NA	X	NA	Bone	4/2/1 Black Rocks Cresent BR4 Box 2a 1 (Bird)	Black Rocks, Wairarapa, Wellington	Archaeological	733 – 566 Cal years BP
AD396	X	X	X	NA	X	NA	X	NA	Bone	3/1/8 Black Rocks Black Midden BR3 N168-9/77	Black Rocks, Wairarapa, Wellington	Archaeological	733 – 566 Cal years BP

Supplementary Table 36. Sequences obtained from GenBank. Sample number refers to tissue samples of which COI was obtained in Cole *et al*. (2018; *Chapter 5*) and which CR has been obtained in this study. Note additional COI sequences were obtained in Cole *et al*. (2018) (*Chapter 5*) but were not used in this study.

Taxon	Accession number	Sequence	Sample number
Aptenodytes forsteri	EU525299	COI	NA
Aptenodytes patagonicus	EU525303	COI	NA
Diomedia exulans	DQ137168	COI	NA
Eudyptes chrysocome	AB775513	CR	NA
Eudyptes chrysocome	DQ525800	COI	NA
Eudyptes chrysolophus chrysolophus	MF469857	COI	NA
Eudyptes chrysolophus chrysolophus	MF469863	COI	NA
Eudyptes chrysolophus chrysolophus	MF469859	COI	NA
Eudyptes chrysolophus chrysolophus	MF469864	COI	NA
Eudyptes chrysolophus chrysolophus	FJ582595	COI	NA
Eudyptes filholi	DQ525795	COI	NA
Eudyptes moseleyi	DQ525790	COI	NA
Eudyptes pachyrhynchus	MF469874	COI	AD257
Eudyptes pachyrhynchus	MF469869	COI	AD259
Eudyptes pachyrhynchus	MF469870	COI	AD260
Eudyptes pachyrhynchus	MF469873	COI	AD261
Eudyptes pachyrhynchus	MF469876	COI	AD262
Eudyptes pachyrhynchus	MF469881	COI	AD264
Eudyptes pachyrhynchus	MF469879	COI	AD265
Eudyptes pachyrhynchus	MF469877	COI	AD266
Eudyptes pachyrhynchus	MF469880	COI	AD267
Eudyptes pachyrhynchus	EU525345	COI	NA
Eudyptes robustus	MF469856	COI	AD270
Eudyptes robustus	MF469855	COI	AD272
Eudyptes robustus	EU525347	COI	NA
Eudyptes chrysolophus schlegeli	MF469865	COI	NA
Eudyptes chrysolophus schlegeli	MF469862	COI	NA
Eudyptes chrysolophus schlegeli	MF469861	COI	NA
Eudyptes chrysolophus schlegeli	MF469858	COI	NA
Eudyptes chrysolophus schlegeli	FJ582599	COI	NA
Eudyptes sclateri	MF469866	COI	AD304
Eudyptula minor	EU525402	COI	NA
Eudyptula novaehollandiae	EU525396	COI	NA
Megadyptes antipodes antipodes	DQ137184	COI	NA
Megadyptes antipodes antipodes	FJ822140	CR	NA
Megadyptes antipodes antipodes	FJ391944	CR	NA
Megadyptes antipodes antipodes	FJ391945	CR	NA
Megadyptes antipodes antipodes	FJ391946	CR	NA
Megadyptes antipodes antipodes	FJ391947	CR	NA
Megadyptes antipodes antipodes	FJ391948	CR	NA
Megadyptes antipodes antipodes	FJ391949	CR	NA
Megadyptes antipodes antipodes	FJ391950	CR	NA
Megadyptes antipodes antipodes	FJ391951	CR	NA
Megadyptes antipodes antipodes	FJ391952	CR	NA
Megadyptes antipodes antipodes	FJ391953	CR	NA

Megadyptes antipodes antipodes	FJ822137	CR	NA
Megadyptes antipodes antipodes	FJ822138	CR	NA
Megadyptes antipodes antipodes	FJ822139	CR	NA
Megadyptes antipodes antipodes	FJ822141	CR	NA
Megadyptes antipodes antipodes	FJ822142	CR	NA
Megadyptes antipodes antipodes	FJ822143	CR	NA
Megadyptes antipodes waitaha	KP644153	CR	NA
Megadyptes antipodes waitaha	FJ391953	CR	NA
Megadyptes antipodes waitaha	FJ391954	CR	NA
Megadyptes antipodes waitaha	FJ391955	CR	NA
Megadyptes antipodes waitaha	FJ391956	CR	NA
Megadyptes antipodes waitaha	FJ391957	CR	NA
Megadyptes antipodes waitaha	FJ391958	CR	NA
Megadyptes antipodes waitaha	FJ391959	CR	NA
Megadyptes antipodes waitaha	FJ391960	CR	NA
Megadyptes antipodes waitaha	FJ391961	CR	NA
Megadyptes antipodes waitaha	FJ391962	CR	NA
Megadyptes antipodes waitaha	FJ391963	CR	NA
Megadyptes antipodes waitaha	FJ391964	CR	NA
Megadyptes antipodes waitaha	FJ391965	CR	NA
Megadyptes antipodes waitaha	FJ391966	CR	NA
Pygoscelis adeliae	DQ137183	COI	NA
Pygoscelis antarctica	EU525474	COI	NA
Pygoscelis papua	EU525486	COI	NA
Spheniscus demersus	EU525504	COI	NA
Spheniscus humboldti	DQ137180	COI	NA
Spheniscus magellanicus	EU525505	COI	NA
Spheniscus mendiculus	DQ137179	COI	NA

Supplementary Table 37. Prior and posterior distributions of parameters for a model of constant population size for *Eudyptes pachyrhynchus* as inferred by DIYABC. Timing of events in the case of this model is irrelevant because N_{ef} remained constant through time. However, these events were kept in the model for consistency with the alternative models that included population declines. Timing of events is in number of generations, assuming a generation time of five years (Otley *et al*., 2017). * indicates that the same priors were used for pre-LGM and pre-human N_{ef} in alternative models tested.

Parameter	Prior distribution	Posterior mode	5 %	95 %
$N_{ef\text{-pre-LGM}}$ and $N_{ef\text{-pre-human}}$ *	Uniform [10 – 10^6]	NA	NA	NA
N_{ef}	Uniform [10 – 10^6]	$5.81X10^4$	$4.28X10^4$	$8.38X10^5$
$T_{\text{pre-human}}$	Uniform [2 – 160]	$1.45X10^2$	$1.23X10^1$	$1.54X10^2$
$T_{\text{pre-LGM}}$	Uniform [2000 – 3000]	$2.87X10^3$	$2.06X10^3$	$2.96X10^3$
substitution rate (COI)	Uniform [10^{-9} – $5X10^{-7}$]	$4.57X10^{-8}$	$3.10X10^{-8}$	$4.45X10^{-7}$

Supplementary Table 38. Comparison of scenarios using the logistic regression approach on the closest 1%, 0.1% and 0.01% of data sets to the observed data, as inferred by DIYABC. See also *Supplementary Figure 24*.

Number of simulations retained	Scenario 1	Scenario 2	Scenario 3	Scenario 4
40000 (1%)	**0.6719** [0.6587,0.6851]	**0.1020** [0.0938,0.1102]	**0.1357** [0.1259,0.1454]	**0.0904** [0.0822,0.0985]
4000 (0.1%)	0.7150 [0.6897,0.7404]	0.0925 [0.0775,0.1075]	0.1341 [0.1142,0.1540]	0.0583 [0.0456,0.0710]
400 (0.01%)	0.6307 [0.2877,0.9738]	0.2188 [0.0000,0.5502]	0.1101 [0.0000,0.3129]	0.0404 [0.0000,0.1695]

Chapter 5

Ancient DNA reveals that the 'extinct' Hunter Island penguin (*Tasidyptes hunteri*) is not a distinct taxon

Publication details, contributions and acknowledgements

Parts of this chapter are published in:

Theresa L Cole[1,2], Jonathan M Waters[1], Lara D Shepherd[3], Nicolas J Rawlence[1], Leo Joseph[4] & Jamie R Wood[2]. (2018). Ancient DNA reveals that the 'extinct' Hunter Island penguin (*Tasidyptes hunteri*) is not a distinct taxon. *Zoological Journal of the Linnean Society.* 182(2): 459 – 464.

Author affiliations:

[1]Department of Zoology, University of Otago, PO Box 56, Dunedin 9054, Otago, New Zealand. [2]Manaaki Whenua Landcare Research, PO Box 69040, Lincoln 7640, Canterbury, New Zealand. [3]Museum of New Zealand Te Papa Tongarewa, PO Box 467, Wellington 6011, Wellington, New Zealand. [4]Australian National Wildlife Collection, CSIRO National Research Collections Australia, GPO Box 1700, Canberra 2601, Australian Capital Territory, Australia.

Author contributions:

Theresa L Cole (20%), Jamie R Wood (20%), Nicolas J Rawlence (20%), Jonathan M Waters (20%) and Leo Joseph (20%) conceived and designed the study. Nicolas J Rawlence (60%), Theresa L Cole (20%), Leo Joseph (15%) and Lara D Shepherd (5%) sampled specimens from museums, Theresa L Cole (80%), Nicolas J Rawlence (10%), Lara D Sheperd (5%) and Jamie R Wood (5%) undertook laboratory work. Theresa L Cole (95%) and Jamie R Wood (5%) analysed the data. Theresa L Cole (30%), Jamie R Wood (30%), Jonathan M Waters (30%) and Lara D Shepherd (10%) contributed to sequencing and laboratory costs. All co-authors helped write the manuscript.

Acknowledgements:

We thank Alan JD Tennyson and Paul Scofield for providing tissue samples, Robert Palmer for arranging loans of specimens from that collection and Trevor Worthy for providing comments on taxonomy. We thank several anonymous reviewers who improved the manuscript. Theresa L Cole was supported by an Otago University Postgraduate Scholarship and the Otago Ecology Fund.

Abstract

The penguin species *Tasidyptes hunteri* van Tets & O'Connor, 1983, the sole representative of an extinct penguin genus, was described on the basis of four bones excavated from a prehistoric midden on Tasmania's Hunter Island. Several authors have since questioned the validity of *T. hunteri*, citing the fragmentary nature of the remains and the similarity of some elements (coracoid and tarsometatarsus) to those of extant *Eudyptes* species. We designed four

overlapping primer pairs to amplify a 499 bp region of COI in penguins and used these primers to amplify and sequence COI from all known bones attributed to *T. hunteri*. The *T. hunteri* COI sequences were assessed within a phylogenetic framework against COI sequences for all extant penguin species. Our results reveal that the *T. hunteri* bones are an assemblage of remains from three extant penguin species (*Eudyptes pachyrhynchus*, *E. robustus*, *Eudyptula novaehollandiae*), and we find no molecular support for any of these bones representing an extinct penguin lineage.

Introduction

Penguins (Sphenisciformes) are a group of flightless marine birds widely distributed throughout the Southern Hemisphere (see *Figure 11A*). The exact number of extant penguin species reported often varies, mainly due to different taxonomic treatment of species and subspecies splits. Nevertheless, based on morphological and phylogenetic analyses (Bertelli & Giannini, 2005; Cole *et al*., 2019b; *Chapter 2*), extant penguins fall into six clearly defined genera; *Aptenodytes*, *Eudyptes*, *Pygoscelis*, *Spheniscus*, *Megadyptes* and *Eudyptula*.

Within some extant penguin species there is clear evidence for loss of genetic lineages as a result of human hunting (e.g. *Eudyptula minor*; see Grosser *et al*., 2015). Moreover, four penguin taxa are thought to have been hunted to extinction during the Holocene. Three of these were cryptic penguin taxa, revealed by aDNA analysis of archaeological midden deposits and natural sub-fossils in NZ and the Chatham Islands: *Megadyptes antipodes waitaha* (Boessenkool *et al*., 2008), *M. a. richdalei* (Cole *et al*., 2019b; *Chapter 2*) and *Eudyptes warhami* (Cole *et al*., 2019b, *Chapter 2*). These taxa were apparently hunted to extinction shortly after Polynesian arrival in the late 13th Century AD (Rawlence *et al*., 2015a; Cole *et al*., 2019b; *Chapter 2*).

The fourth example is *Tasidyptes hunteri*; van Tets and O'Connor, 1983. The description of this species was based on morphological analysis of bones from a prehistoric midden site on Hunter Island, just north of Tasmania. Available material of *T. hunteri* includes a complete pelvis in three parts (ANWC B21141, originally published as ANWC BS2670; holotype), a tarsometatarsus (ANWC B21143, BS2668; paratype), a corocoid (ANWC B21142, BS2669; paratype) and a synsacrum (ANWC B22669, BS2667; paratype) (see *Figure 28* and *Supplementary Table 39*). All type specimens were excavated from amongst archaeological material on Hunter Island and are thought to represent at least two individuals (van Tets & O'Connor, 1983). Based on a single radiocarbon date and correlation with other middens on Hunter Island, the age of the bones are estimated to be <1600 cal years BP (van Tets &

O'Connor, 1983). In addition, two other bones from the ANWC are also referred to *T. hunteri*. These are a coracoid (ANWC B21145, originally registered as BS6618) from Louisa River Cave, Louisa Bay, southern Tasmania, and a coracoid (ANWC B21144, BS6612) from Prion Beach, southern Tasmania (see *Figure 29*).

Figure 28. *Tasidyptes hunteri* holotype and paratypes. **(A)** is the coracoid ANWC B21145 from Louisa Bay; **(B)** is the coracoid ANWC B21144 from Prion Beach; **(C)** is the coracoid ANWC B21142 from Hunter Island; **(D – F)** is the pelvis in three parts ANWC B21141 from Hunter Island; **(G)** is the tarsometatarsus ANWC B21143 from Hunter Island; and **(H)** is the synsacrum ANWC B22669 from Hunter Island. Scale bar is 50 mm each block. Photographs: Jamie Wood.

In describing *T. hunteri*, van Tets and O'Connor (1983) compared morphological characteristics of the Hunter Island bones to those of 16 fossil and extant penguin taxa. They claimed that the *T. hunteri* bones were morphologically distinct, yet most closely resembled those of *Megadyptes*, *Eudyptes* and *Eudyptula*. The diagnosis of *T. hunteri* considered that the caudal part of the synsacrum differed from the three above-mentioned genera in having relatively broader fused vertebrae and longer, more slender lateral processes. Moreover, the lateral foramen vasculare proximale was situated more distally than the medial foramen vasculare proximale on the plantar surface of the tarsometatarsus (van Tets & O'Connor, 1983; Park & Fitzgerald, 2012). The distinctiveness of *T. hunteri*, however, has since been questioned by several authors (e.g. Fordyce & Jones, 1990; Ksepka & Clarke, 2010; Park & Fitzgerald, 2012) on multiple grounds. First, the remains are fragmentary (e.g. the holotype was found in three separate parts). Second, the coracoid and the tarsometatarsus are indistinguishable from those of *Eudyptes* species. Third, the four type specimens were excavated from different stratigraphic horizons of the archaeological site and so may not represent a single taxon (Park & Fitzgerald, 2012).

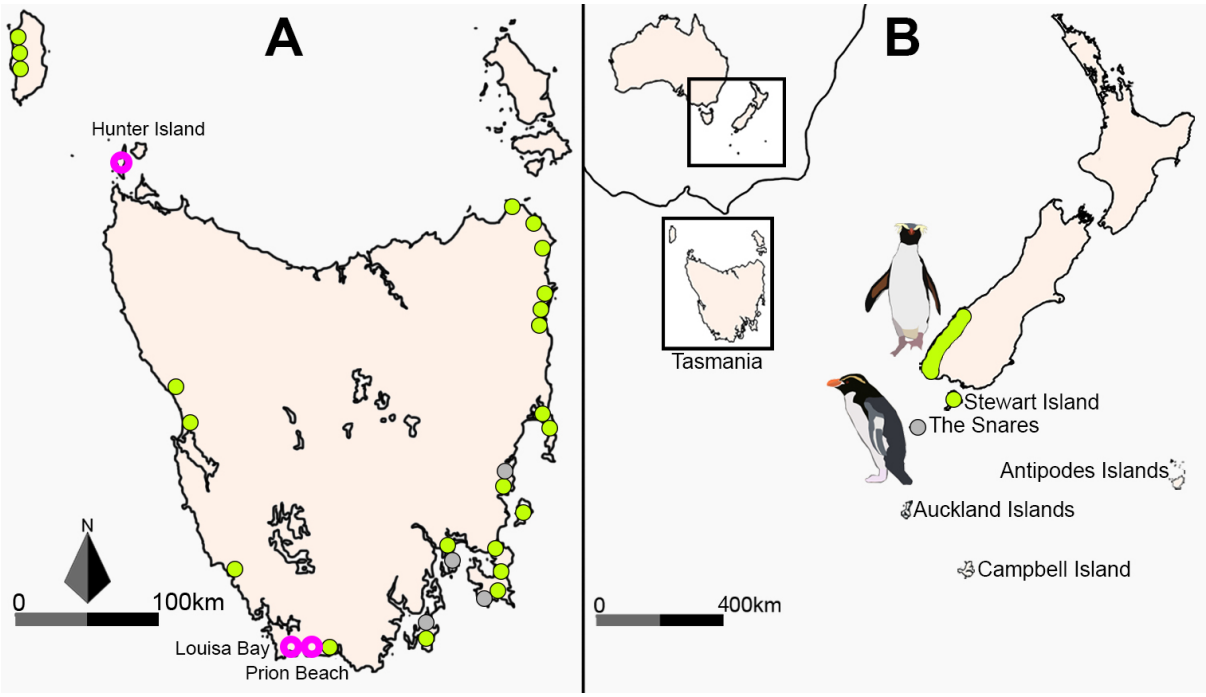


Figure 29. Tasmanian sites where *Tasidyptes hunteri* and vagrant *Eudyptes pachyrhynchus* and *E. robustus* have been recorded. **(A)** *Tasidyptes hunteri* locations are marked in pink, *Eudyptes pachyrhynchus* is marked in green and *E. robustus* is marked in grey. These vagrant records were recorded between 1970 – 1991. *Eudyptula novaehollandiae* are not shown as they breed across much of coastal Tasmania. **(B)** Map of breeding localities of the endemic *Eudyptes pachyrhynchus* and *E. robustus* on NZ and The Snares Islands.

Unfortunately, discrimination of *Eudyptes* species, and even between *Eudyptes* and *Megadyptes*, based on post-cranial bones is challenging given the high osteological similarities between species and extreme sexual dimorphism within species (see also Cole *et al*., 2019b; *Chapter 2*). This problem is further exacerbated when skeletal remains are fragmentary. Worthy (1997), for example, highlighted challenges in identifying differences between post-cranial elements of *Eudyptes pachyrhynchus*, *E. robustus*, *E. sclateri* and *Megadyptes antipodes*. Worthy (1997) found that many osteological features of these taxa to be very similar, noting that elements overlapped in size between species in these two genera.

Molecular analyses offer an alternative approach for elucidating the taxonomic relationships of morphologically similar archaeological remains. Here, we present aDNA analyses of all archaeological remains attributed to *Tasidyptes hunteri* in order to assess the phylogenetic status of this taxon. In doing so we also present new penguin specific primers designed to amplify a diagnostic 499 bp fragment of the COI gene, a marker which has been widely used to delineate bird species (Tizard, 2019) and, in some cases, identify cryptic species (Hebert *et al*., 2004). As such, these primers were subsequently used to identify between *Eudyptes* and *Megadyptes* taxa in Cole *et al*. (2019b; *Chapter 2*) and Cole *et al*. (2019a; *Chapter 4*).

Materials and Methods

DNA Extraction

Ancient DNA extractions were performed in dedicated aDNA facilities at the Department of Zoology, University of Otago, Dunedin, MWLR, Lincoln, and the NMNZ, Wellington. Standard aDNA protocols, as advocated by Cooper and Poinar (2000) were used, including the use of physically isolated dedicated aDNA facilities and extraction and PCR controls. DNA was extracted from each specimen using the Qiagen DNeasy® Tissue Extraction Kit (Qiagen Inc., Chatsworth, California, USA) following a modified protocol adapted for bone collagen (Thomson *et al*., 2014).

Primer design, polymerase chain reaction and sequencing

We designed penguin specific internal primers (*Supplementary Table 40*) to amplify a 499 bp diagnostic region of COI in four short overlapping fragments, suitable for degraded or aDNA. Primers were tested on 22 historical tissue samples from all *Eudyptes* taxa except *E. moseleyi* and *E. chrysocome* (*Supplementary Tables 41*). The primers were also used to amplify COI from each *Tasidyptes hunteri* specimen (*Supplementary Table 40*), with each of the four fragments being amplified and sequenced at least twice for each specimen. PCRs (12.5 μL each) were performed using 2 mg/mL RSA (Sigma), 1 X PCR buffer, 2 mM $MgSO_4$, 80 μM dNTPs, 0.4 μM each primer, 0.625 U of HiFi Platinum Taq (Invitrogen), and 1 μL DNA on a BIO-RAD MyCycler thermal cycler with an initial denaturation of 94°C for 3 min; followed by 55 cycles of 94°C for 30 s, 59°C for 30 s and 68°C for 45 s; and a final extension at 68°C for 10 min. PCR products were purified using SPRIselect (Beckman Coulter, Inc., Indianapolis, IN, USA) and sequenced at the MWLR sequencing facility on an Applied Biosystems 3500xL Genetic Analyzer in Auckland.

Data analysis

We used Geneious R8 (Biomatters) to examine, edit and align forward and reverse consensus sequences for each specimen. We used MEGA v.6 to align sequences with the same region of COI for all penguin taxa obtained from GenBank and the Barcode of Life Database (BOLD) (see *Supplementary Table 42*). A maximum-credibility phylogeny was created using BEAST, with a HKY model (Hasegawa *et al*., 1985) with four gamma categories, Yule-speciation prior, and MCMC chain length of 30 million, recording every 1000 states with a 10% burnin.

Results

We successfully amplified between 296 – 499 bp of the COI gene from all specimens using the four overlapping primer pairs (Genbank accession numbers MF469854 – MF469881). With the exception of *Eudyptes chrysolophus chrysolophus*/*E. c. schlegeli* complex (see *Chapter 3*), our primers successfully differentiated every *Eudyptes* species including the occasionally merged *E. pachyrhynchus*/*E. robustus*) (see also *Chapter 3*). In addition, these primers also amplified sequences effective in distinguishing between *Eudyptes* and *Eudyptula* penguin remains.

Based on mtDNA relationships, the results from our analysis suggest that the bones referred to *Tasidyptes hunteri* are an artificial assemblage comprising three extant penguin species belonging to two genera (*Figure 30*): the holotype, paratype, Hunter Island coracoid, and Louisa River coracoid are phylogenetically attributable to *Eudyptes pachyrhynchus*, the coracoid from Prion Beach is attributable to *E. robustus*, and the synsacrum is attributable to *Eudyptula novaehollandiae*.

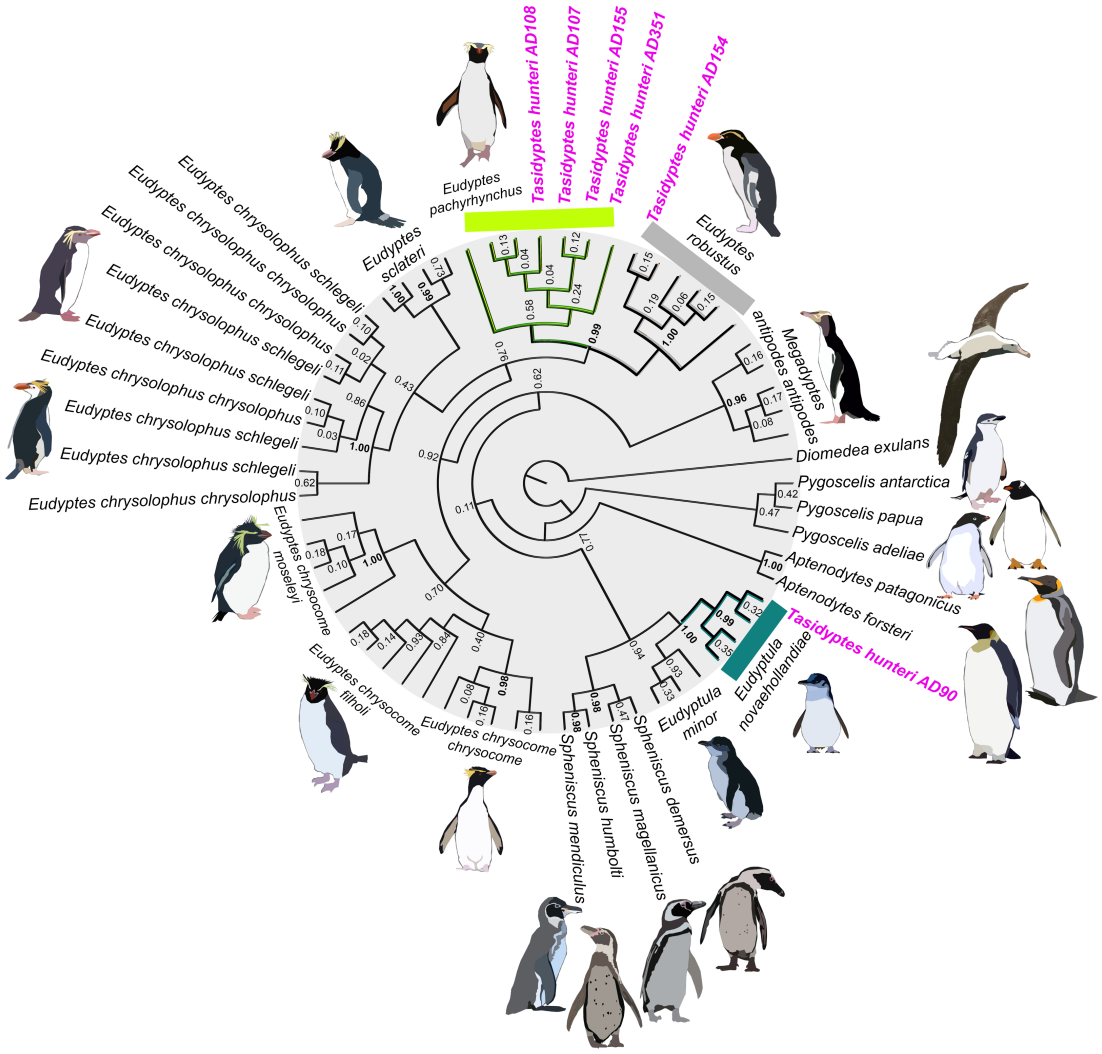


Figure 30. Bayesian phylogenetic tree of penguins, including those attributed to *Tasidyptes hunteri.* Posterior probabilities are shown.

Discussion

Ancient DNA results of the current analysis strongly indicate that *T. hunteri* is not a valid taxon, but instead represents an artificial assemblage of remains from extant penguin taxa. Moreover, where preserved, osteological characters of each bone (following Bertelli & Giannini, 2005) reflected the states present in the species to which it was assigned based on DNA. These findings contradict van Tets and O'Connor's (1983) claim of a Holocene penguin extinction in southern Australia. Based on our analysis of the type material, we place *Tasidyptes hunteri* van Tets and O'Connor, 1983 in synonymy with *Eudyptes pachyrhynchus* Gray 1845.

The presence of three species of penguin (formerly attributed to *Tasidyptes*) in Tasmania's archaeological record is not surprising given the distributions and movement patterns of these taxa within the Australasian region. For example, *Eudyptula* is distributed along the southern coastline of Australia including Tasmania, Bass Strait Islands and NZ. Recent research has revealed that *Eudyptula* comprises two distinct lineages accorded species status by Grosser *et al*. (2015): *Eudyptula novaehollandiae*, historically endemic to Australia but now also on the Otago coast of NZ, and *Eudyptula minor*, endemic to NZ. Our results reveal that the synsacrum belonged to the primarily Australian *E. novaehollandiae*. It is likely that this bone was from a local colony, given that other *Eudyptula* bones were also recorded in the Hunter Island site (van Tets & O'Connor, 1983). Although other penguin species do not breed in Australia, at least eight have been recorded as vagrants in Tasmania (Iredale & Cayley, 1925; Hindwood, 1938; Woehler, 1992; Simpson, 2008; Eric Woehler, personal communication). Of these, *Eudyptes robustus* has regularly been recorded in the Southern Ocean waters near Australia during winter, around 5000 km from their breeding colonies on The Snares and Western Chain (Woehler, 1992; Simpson, 2008; David Thompson, personal communication). A similar winter movement pattern is exhibited by *E. pachyrhynchus*, where birds frequently disperse to southern Australian waters from their natal grounds along the southern NZ coast (Warham, 1974b; Mattern *et al*., 2018; see *Figure 31*). These dispersal patterns may explain both the relatively frequent occurrence of these species in Tasmania and southern mainland Australia (44 records of *E. pachyrhynchus*, and five of *E. robustus* in Tasmania between 1970 – 1991; Woehler, 1992; see *Figure 29*), and their presence in archaeological deposits from this region.

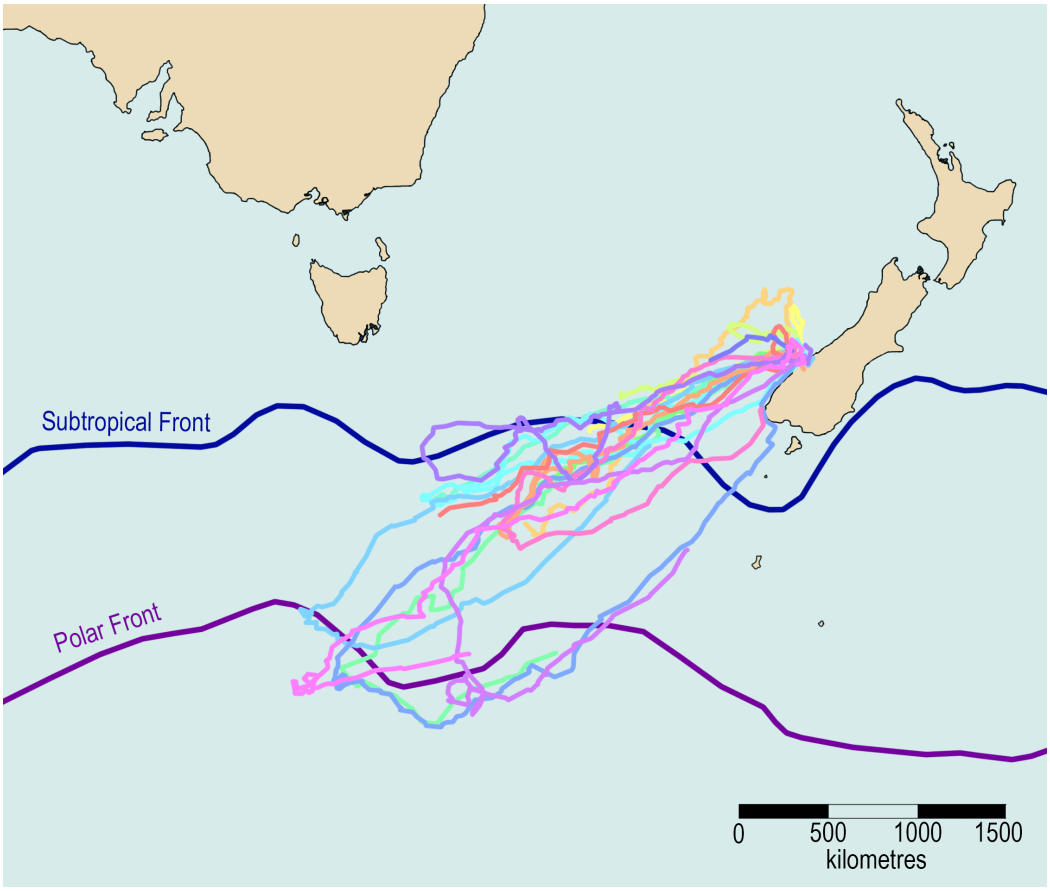


Figure 31. Travelling paths of 17 *Eudyptes pachyrhynchus* individuals during their pre-moult journey between November 2016 and March 2017 based on satellite tags. Each coloured line represents a different individual. Note that data for only five complete paths were collected. This figure is adapted from Mattern *et al*. (2018).

Nevertheless, our finding that the Hunter Island penguin bones include two vagrant species is not without precedent. Vagrant species, particularly medium to large sized birds, have been recorded previously from archaeological and prehistoric midden deposits around the world. For example, bones from at least ten individuals of the Australian pelican (*Pelecanus conspicillatus*) known from eight prehistoric midden deposits throughout NZ are considered to represent vagrant individuals (Lalas *et al*., 2014). A single vagrant sub-Antarctic *Megadyptes antipodes antipodes* has also been identified in a prehistoric midden deposit from NZ prior to the extinction of *M. a. waitaha* (Boessenkool *et al*., 2008) and vagrant *E. warhami* specimens have been identified from archaeological deposits on mainland NZ (Cole *et al*., 2019a; *Chapter 4*). It is thought that cinereous vulture (*Aegypius monachus*) remains from the Roman period of the Netherlands and Belgium may also represent vagrant individuals (Groot *et al*., 2010). We suggest that there could even be a bias towards over-representation of large vagrant bird species in midden deposits. Such individuals may have attracted the attention of hunters due to being out of the ordinary, and may have temporarily been more docile and hence easier to catch due to the state of exhaustion commonly exhibited by long-distance vagrants (e.g. Bried, 2003; Lees & Gilroy, 2009).

Conclusions

Mitochondrial DNA sequences reveal that the assemblage of bones upon which *Tasidyptes hunteri* was described includes more than one taxon (supporting the suggestion of Park and Fitzgerald, 2012). In fact, the assemblage represents three extant species: *Eudyptula novaehollandiae*, *Eudyptes pachyrhynchus* and *E. robustus*. The holotype of *Tasidyptes hunteri* was demonstrated to be from *Eudyptes pachyrhynchus*, and so *Tasidyptes hunteri* van Tets and O'Connor, 1983 is placed in the synonymy of *Eudyptes pachyrhynchus* Gray 1845. *Eudyptula novaehollandiae* remains the only penguin species known to have bred around mainland Australia during the Holocene. Our study provides a further example of how aDNA analyses can help resolve issues around the taxonomic identity of prehistoric remains.

Supplementary Tables for *Chapter 5*

Supplementary Table 39. *Tasidyptes hunteri* specimens.

Sample Code	Museum Number	Original publication code in van Tets and O'Connor 1983	Sampling Locality	Element	Taxon	Sequence length (bp)	Genbank accession number
AD90	ANWC B22669	ANWC BS2667	Hunter Island	synsacrum	*Eudyptula novaehollandiae*	296	MF469878
AD107	ANWC B21142	ANWC BS2669	Hunter Island	coracoid	*Eudyptes pachyrhynchus*	356	MF469868
AD108	ANWC B21143	ANWC BS2668	Hunter Island	tarsometatarsus	*Eudyptes pachyrhynchus*	418	MF469871
AD154	ANWC B21144	ANWC BS6612	Prion Beach	coracoid	*Eudyptes robustus*	499	MF469854
AD155	ANWC B21145	ANWC BS6618	Louisa Bay	coracoid	*Eudyptes pachyrhynchus*	499	MF469875
AD351	ANWC B21141	ANWC BS2670	Hunter Island	pelvis	*Eudyptes pachyrhynchus*	455	MF469872

Supplementary Table 40. Primers designed for amplifying COI in penguins. Amplicon length excludes primer sequences.

Primer name	Primer sequence (5' – 3')	3' position on *Eudyptes chrysocome*	Amplicon
Eud_COI_1	GNGACGACCAAATCTACAACGTAA	5666	142
Eud_COI_1a	GTAGGAAGGAAGGGGGGAGT	5809	
Eud_COI_2	GACATAGCATTCCCCCGCATG	5788	140
Eud_COI_2a	GGTGGAGTGAGAARATRGCTAAG	5929	
Eud_COI_3	GCTGGYACAGGATGRACTGTA	5881	141
Eud_COI_3a	RGGGGTTTGRTACTGTGAGAG	6023	
Eud_COI_4	ATTAACTTCATCACCACCGCC	6001	164
Eud_COI_4a	ATTGGGTCACCTCCTCCG	6166	

Supplementary Table 41. COI primer success across *Eudyptes* samples. X indicates that the COI region did not amplify.

Sample Number	Museum Number	Sampling Locality	EudCOI1 and EudCOI1a	EudCOI2 and EudCOI2a	EudCOI3 and EudCOI3a	EudCOI4 and EudCOI4a	Taxon	Genbank accession number
AD270	OR.019353	The Snares	Y	Y	Y	Y	*Eudyptes robustus*	MF469856
AD272	OR.026830	Penguin Creek, The Snares	Y	Y	Y	Y	*Eudyptes robustus*	MF469855
AD257	OR.026082	Dusky Sound	Y	Y	Y	Y	*Eudyptes pachyrhynchus*	MF469874
AD259	OR.015371	Jackson Head, Westland	Y	Y	Y	N	*Eudyptes pachyrhynchus*	MF469869
AD260	OR.015368	Jackson Head, Westland	Y	Y	Y	Y	*Eudyptes pachyrhynchus*	MF469870
AD261	OR.023566	Turnbull River Mouth, Westland	Y	Y	Y	Y	*Eudyptes pachyrhynchus*	MF469873
AD262	OR.026575	Halfmoon Bay, Stewart Island	Y	Y	Y	Y	*Eudyptes pachyrhynchus*	MF469876
AD264	OR.015835	Jackson Head, Westland	N	Y	Y	Y	*Eudyptes pachyrhynchus*	MF469881
AD265	OR.015370	Jackson Head, Westland	N	Y	N	Y	*Eudyptes pachyrhynchus*	MF469879
AD266	OR.001067	Solander Island	Y	Y	Y	Y	*Eudyptes pachyrhynchus*	MF469877
AD267	OR.015369	Jackson Head, Westland	N	Y	Y	Y	*Eudyptes pachyrhynchus*	MF469880
AD304	OR.021942	Antipodes Island	Y	Y	Y	N	*Eudyptes sclateri*	MF469866
AD175	OR.012822	South Thule Island, South Sandwich Islands	Y	Y	Y	N	*Eudyptes chrysolophus chrysolophus*	MF469857
AD176	OR.013284	Macquarie Island	Y	Y	Y	Y	*Eudyptes chrysolophus chrysolophus*	MF469865
AD178	OR.013282	Macquarie Island	Y	Y	Y	N	*Eudyptes chrysolophus chrysolophus*	MF469861
AD179	OR.010287	Cape Hallett, Antarctica	Y	Y	Y	N	*Eudyptes chrysolophus chrysolophus*	MF469863
AD180	OR.012821	South Thule Island, South Sandwich Islands	Y	Y	Y	N	*Eudyptes chrysolophus chrysolophus*	MF469859
AD181	OR.010286	Sabrina Island	Y	Y	Y	N	*Eudyptes chrysolophus chrysolophus*	MF469864
AD177	OR.013285	Macquarie Island	Y	Y	Y	N	*Eudyptes chrysolophus schlegeli*	MF469862
AD182	OR.013286	Macquarie Island	Y	N	Y	N	*Eudyptes chrysolophus schlegeli*	MF469858
AD208	OR.000702	Macquarie Island	N	Y	Y	N	*Eudyptes chrysolophus schlegeli*	MF469860
AD253	Av.21761	Campbell Island	Y	N	Y	N	*Eudyptes filholi*	MF469867
Total successful reactions per primer pair			**18**	**20**	**21**	**11**		

Supplementary Table 42. Sequences obtained from GenBank and BOLD.

Taxon	Database	Accession number
Eudyptes pachyrhynchus	Genbank	DQ137170
Eudyptes pachyrhynchus	Genbank	EU525344
Eudyptes pachyrhynchus	Genbank	EU525345
Eudyptes robustus	Genbank	DQ137176
Eudyptes robustus	Genbank	EU525346
Eudyptes robustus	Genbank	EU525347
Eudyptes robustus	Genbank	EU525348
Eudyptes robustus	Genbank	EU525349
Eudyptes sclateri	Genbank	DQ137169
Eudyptes sclateri	BOLD	GBIR0I83
Eudyptes sclateri	BOLD	BROMB075
Eudyptes sclateri	BOLD	BROMB076
Eudyptes chrysolophus schlegeli	Genbank	FJ582596
Eudyptes chrysolophus schlegeli	Genbank	FJ582597
Eudyptes chrysolophus schlegeli	Genbank	FJ582598
Eudyptes chrysolophus schlegeli	Genbank	FJ582599
Eudyptes chrysolophus schlegeli	Genbank	DQ137175
Eudyptes chrysolophus chrysolophus	Genbank	DQ137171
Eudyptes chrysolophus chrysolophus	Genbank	FJ582593
Eudyptes chrysolophus chrysolophus	Genbank	FJ582594
Eudyptes chrysolophus chrysolophus	Genbank	FJ582595
Eudyptes moseleyi	Genbank	DQ525786
Eudyptes moseleyi	Genbank	DQ525787
Eudyptes moseleyi	Genbank	DQ525788
Eudyptes moseleyi	Genbank	DQ525789
Eudyptes moseleyi	Genbank	DQ525790
Eudyptes filholi	Genbank	DQ525791
Eudyptes filholi	Genbank	DQ525792
Eudyptes filholi	Genbank	DQ525793
Eudyptes filholi	Genbank	DQ525794
Eudyptes filholi	Genbank	DQ525795
Eudyptes chrysocome	Genbank	DQ525796
Eudyptes chrysocome	Genbank	DQ525797
Eudyptes chrysocome	Genbank	DQ525798
Eudyptes chrysocome	Genbank	DQ525799
Eudyptes chrysocome	Genbank	DQ525800
Eudyptula minor	Genbank	EU525399
Eudyptula minor	Genbank	EU525400
Eudyptula minor	Genbank	EU525401
Eudyptula novaehollandiae	Genbank	EU525393
Eudyptula novaehollandiae	Genbank	EU525394
Eudyptula novaehollandiae	Genbank	EU525395
Megadyptes antipodes	Genbank	DQ137184
Megadyptes antipodes	Genbank	AY369062
Megadyptes antipodes	BOLD	BROMB027
Megadyptes antipodes	BOLD	BROMB046
Megadyptes antipodes	BOLD	BROMB047
Spheniscus mendiculus	Genbank	DQ137179
Spheniscus humbolti	Genbank	DQ137180
Spheniscus magellanicos	Genbank	EU525505
Spheniscus demersus	Genbank	EU525504
Aptenodyptes forsteri	Genbank	EU525299
Aptenodyptes patagonicus	Genbank	EU525303
Pygoscelis adeliae	Genbank	DQ137183
Pygoscelis papua	Genbank	EU525486
Pygoscelis antarctica	Genbank	EU525474

Chapter 6

Synthesis, recommendations and conclusions

Synthesis

The aim of this thesis was to understand how geological, climatic and human settlement histories have contributed to the evolution and biodiversity of penguin assemblages on Southern Ocean islands. As penguins represent a biologically diverse and widespread marine bird radiation, they provide a fascinating system for examining abiotic and biotic drivers of evolution.

This thesis used a variety of approaches to test several broad evolutionary hypotheses. These included short DNA sequences (COI, CR), genome-wide markers (mitogenomes, SNPs) and morphological data (skeletal characters), analysed together under phylogenetic and demographic frameworks. Collectively, this thesis used penguins as a model system to demonstrate that (1) the recent emergence of islands in the Southern Ocean has promoted endemic species diversification; (2) climate change following the LGM promoted rapid circumpolar recolonisation to sub-Antarctic islands; (3) Polynesian arrival to NZ and the Chatham Islands resulted in extinction of island-endemic taxa, and (4) island geography and habitat preferences buffered several species from human-driven extirpation. Taken together, results from this thesis provide compelling evidence for the effects of geological, climatic and anthropogenic impacts on Southern Ocean marine biodiversity. In this final chapter, I expand and connect the broad themes of the previous chapters, and place my research within the broader context of island biology.

Major findings of this study

Systematics and phylogenetics of modern penguins

Phylogenetic results presented in this thesis support previous studies in recovering *Spheniscus* + *Eudyptula* and *Eudyptes* + *Megadyptes* clades (Bertelli & Gianni, 2005; Baker *et al*., 2006; Ksepka *et al*., 2006). *Aptenodytes* + *Pygoscelis* were also recovered as basal to all other extant penguins, supporting recent findings of Subramanian *et al*. (2013) and Gavryushkina *et al*. (2017), despite other studies having recovered *Aptenodytes* alone as basal to all other extant penguins (e.g. Bertelli & Gianni, 2005; Baker *et al*., 2006; Ksepka *et al*., 2006), including a

recent study using high-coverage full mitogenome and nuclear genomes (Pan *et al*., in press). As such, it is almost certain that *Aptenodytes* is basal, and sister to all extant penguins. Nevertheless, phylogenetic results in this thesis have nonetheless provided new insights into the systematics of penguins, especially for the penguin genera *Eudyptes*, *Megadyptes*, *Eudyptula* and *Tasidyptes*. The following section provides the most up-to-date systematic recommendations, based on new molecular and morphological data from this thesis.

Eudyptes

Ancient DNA analysis of sub-fossil bones, combined with detailed morphological comparisons, led to the formal taxonomic description of *Eudyptes warhami* from the Chatham Islands archipelago. Although *E. warhami* was also detected in low numbers on the NZ mainland (referred to as *Eudyptes* clade X in Cole *et al*., 2019a; *Chapter 4*), the relatively large number of sub-fossil bones of this lineage present on the Chatham Islands, compared to the few found on the mainland, suggest that *E. warhami* was an endemic breeder on the Chatham Islands. The presence of *E. warhami* on the NZ mainland may reflect foraging or over-wintering ranges. The formal taxonomic description of this species was achieved through a combination of morphological and molecular analyses. Phylogenetic results revealed that *E. warhami* was sister to the extant Antipodes and Bounty Islands-endemic *E. sclateri*, which is commonly observed as a vagrant in both mainland NZ and the Chatham Islands (Powlesland, 1984; Marchant & Higgins, 1990; Taylor, 2000; Robertson *et al*., 2017; Mattern & Wilson, 2018). While previous studies (e.g. Tennyson & Millener, 1994) had suggested the presence of an extinct *Eudyptes* penguin on the Chatham Islands, no previous study had made a formal description of this unique penguin lineage.

E. pachyrhynchus and *E. robustus* have only recently been recognised as separate species by Birdlife International (Mattern & Wilson, 2018). In the 1970s the Checklist of NZ Birds considered *E. pachyrhynchus*, *E. robustus* and *E. sclateri* as conspecific (Kinsky, 1970). Species-level splits for these lineages were favoured by Stonehouse (1971), Falla *et al*. (1974) and Warham (1974b), based on distinct morphological and ecological features. More recently, Davis and Renner (2003), Baker *et al*. (2006) and Ksepka *et al*. (2006) examined the evolutionary relationships of these taxa using morphology, protein data and short DNA sequences, supporting the separation of three species. However, a subsequent taxonomic review by Christidis and Boles (2008) still considered *E. pachyrhynchus* and *E. robustus* as conspecific. COI, CR, mitogenomes and genome-wide SNPs presented in this thesis all support species-level distinctions of *E. pachyrhynchus*, *E. robustus* and *E. sclateri*.

Previous research has suggested that populations of *E. robustus* breeding on The Snares and the Western Chain might represent separate taxa, based on slight differentiation in morphology and reproductive timing (Miskelly, 1997). COI, CR and genome-wide SNPs analysed here demonstrate no separation, and indicate that the two populations should remain synonymous.

In addition to the taxonomic uncertainty regarding NZ's endemic *Eudyptes* fauna described above, there has been substantial controversy regarding the taxonomy of the rockhopper penguin complex (*E. moseleyi*, *E. chrysocome* and *E. filholi*). In the 1990s, rockhopper penguins were considered to represent a single species (Martínez, 1992). Subsequently, based on genetic and behavioural studies, they were separated into two taxa; *E. moseleyi* and *E. chrysocome* (Jouventin *et al*., 2006). More recently, Banks *et al*. (2006), de Dinechin *et al*. (2009) and Frugone *et al*. (2018) suggested that rockhopper penguins should be separated into three separate species; *E. moseleyi*, *E. chrysocome* and *E. filholi*. Nevertheless, some authors still consider *E. chrysocome* and *E. filholi* as a single species, or consider *E. filholi* as a subspecies of *E. chrysocome* (Dickinson & Remsen, 2013; Pütz *et al*., 2013; Mays *et al*., in press). COI, CR, mitogenomes, and genome-wide SNPs from this thesis support the species-level distinction of three rockhopper penguin taxa.

The taxonomic status of *E. chrysolophus chrysolophus* versus *E. c. schlegeli* has similarly remained controversial. The Macquarie Island-endemic *E. c. schlegeli* is distinguished from the more-widespread *E. c. chrysolophus* by the former's slightly larger size, white face, and localised breeding range. Based on these apparent morphological and ecological differences, the two taxa are commonly considered as distinct species (Bertelli & Giannini, 2005; Ksepka *et al*., 2006; Salton *et al*., 2019), or occasionally as subspecies (Brown & Klages, 1987; Christidis & Boles, 2008). Mitochondrial DNA sequences, mitogenomes, and genome-wide SNPs from this thesis demonstrate that *E. c. chrysolophus* and *E. c. schlegeli* are genetically indistinguishable, with very little population structure within or between these taxa. This finding is further supported by recent analyses of short mitochondrial and nuclear sequences (Frugone *et al*., 2018) and genome-wide SNPs (Frugone *et al*., 2019). While robust genetic evidence strongly suggests that *E. c. chrysolophus* and *E. c. schlegeli* are incipient species (or in the earliest phase of speciation), in the absence of non-molecular evidence, such as differences in behaviour and ecology, they are tentatively referred to as subspecies of *E. chrysolophus*.

Megadyptes

Megadyptes is a species-poor penguin genus, being restricted to NZ waters. Only two *Megadyptes* taxa have previously been described: the extant *M. antipodes*, and the slightly smaller, extinct *M. waitaha* (Boessenkool *et al*., 2008). While *M. antipodes* was once endemic to the NZ sub-Antarctic Campbell and Auckland Islands, the taxon has recently (last 400 years) expanded its range to the oceanic-temperate South Island of NZ, following the extinction of *M. waitaha* (Boessenkool *et al*., 2008; Rawlence *et al*., 2015a), which formerly occurred on NZ's South and North Islands (Boessenkool *et al*., 2008; Rawlence *et al*. 2018). Boessenkool *et al*. (2008) described *M. waitaha* as a full species, based on CR sequences and comparison of 10 morphological characters. Results from this thesis uncovered a third, even smaller, *Megadyptes* taxon from the Chatham Islands. Further aDNA analyses, using the same CR sequences of Boessenkool *et al*. (2008), as well as partial mitogenomes demonstrated that the Chatham Islands taxon was a distinct genetic lineage of *M. antipodes*, even though it was closer in size to *M. waitaha*. A re-analysis of the morphological characters of Boessenkool *et al*. (2008), combined with a comparison of additional morphological characters across all *Megadyptes* taxa, demonstrated that only size can reliably be used to morphologically distinguish any of the *Megadyptes* taxa. Therefore, in contrast to Boessenkool *et al*. (2008), results from this thesis concludes that there is no strong morphological differentiation between *M. antipodes* and *M. waitaha*. For this reason, *Megadyptes* is revised as a monotypic genus, comprising three distinct subspecies, including the newly-described Chatham Islands taxon *M. a. richdalei*.

Eudyptula

A recent genetic and morphological study by Grosser *et al*. (2017) recommended that *Eudyptula* penguins be recognised as two distinct taxa, which comprise the endemic NZ *Eudyptula minor* and the endemic Australian (and recently expanded Otago) *E. novaehollandiae* (Grosser *et al*., 2016). Nevertheless, some authors still consider the two taxa as conspecific (Mattern & Wilson, 2018), arguing that ecological research still needs to be undertaken to recognise the two taxa as distinct species. Analysis of mitogenomes demonstrates a deep evolutionary split between *E. minor* and *E. novaehollandiae* of 3.3 Ma, supporting Grosser *et al*. (2017) for the distinction of two taxa.

Previous studies have also suggested that populations of *E. minor* breeding on the NZ Banks Peninsula represent a separate species or subspecies of *E. minor*, based on slight differences in morphology (Meredith & Sin, 1988a; Meredith & Sin 1988b; Challies & Burleigh, 2004; Baker *et al*. 2006). While this thesis does not explicitly address the status of this population, recent genetic data of Grosser *et al*. (2016) show no evidence that this phenotypically distinct

population (previously referred to as *E. m. albosignata*) represents a genetically distinct lineage of *E. minor*.

Tasidyptes

Hunter Island's extinct penguin *Tasidyptes hunteri* was formally described by van Tets and O'Connor in 1983. Several authors (e.g. Fordyce & Jones, 1990; Ksepka & Clarke, 2010; Park & Fitzgerald, 2012) have since questioned the taxonomic validity of *T. hunteri*, citing the fragmentary nature of the remains and the similarity of some elements to those of extant *Eudyptes* penguins, while other studies have incorporated *T. hunteri* into phylogenetic analyses (Park *et al*., 2016). This thesis genetically tested all *T. hunteri* specimens and demonstrated that the taxon is actually based on an assemblage of sub-fossil bones derived from three different extant penguin species (*Eudyptes robustus*, *E. pachyrhynchus* and *Eudyptula novaehollandiae*). *T. hunteri* therefore, is not a valid taxon.

Evolution and biogeography of modern penguins

Evolution of modern penguins is linked to the emergence of geologically young islands

This thesis provides the first temporal phylogenetic evidence suggesting that the emergence of young islands within the past five million years may have facilitated the evolution of several island-endemic penguin lineages. This thesis incorporated extensive new data (compared to previous phylogenetic studies of Baker *et al*., 2006; Ksepka *et al*., 2006; Subramanian *et al*., 2013; Frugone *et al*., 2018 and Gavryushkina *et al*., 2017; see *Figure 6*), including mitogenomes for four additional penguin taxa, and new temporal calibration approaches derived from the fossil record. Divergence dates from this study imply that island emergence has been key to the evolution of *Spheniscus mendiculus*, *Eudyptes warhami*, *E. moseleyi*, *E. chrysolophus schlegeli* and possibly *Eudyptes robustus* and *Megadyptes antipodes richdalei*.

Previous studies have suggested that the intersection of major ocean currents, such as ocean fronts may have promoted divergence between genetically structured rockhopper penguin species (de Dinechin *et al*., 2009; Frugone *et al*., 2018). For example, de Dinechin *et al*. (2009) and Frugone *et al*. (2018) suggested that the breeding range of rockhopper penguins are geographically separated by the STF. However, *E. moseleyi* (which is considered to only breed on islands north of the STF), actually breeds on islands both north (Tristan da Cunha, Amsterdam and St Paul Islands), and south (Gough Island) of the STF. The alternative hypothesis for the divergence of *E. moseleyi* from *E. chrysocome*/*E. filholi*, as put forward in this thesis, is that the recent volcanic emergence of Gough Island led to the original isolation of

founding *E. moseleyi* individuals (which may have been promoted by the ecological productivity in surrounding waters), followed by dispersal and colonisation of the younger Amsterdam and St Paul Islands. Limited evidence (e.g. shallow genetic structure) suggests that ocean fronts may have partly reduced gene flow between extant populations (discussed below), however, there is no evidence to indicate that ocean fronts have played a major role in the evolution of distinct taxa on a geological timescale.

Based on phylogenetic analyses in this thesis, the divergence of other widespread penguin taxa such as *E. filholi*, *E. chrysocome*, *E. chrysolophus chrysolophus*, *Aptenodytes patagonicus*, *A. forsteri* and *Pygoscelis papua* generally occurred during the glacial and inter-glacial periods of the Pleistocene, pre-dating the islands that those taxa now breed on. While it remains possible that island emergence may have played a role in the diversification of those taxa, fossil records of the first diverging lineages are likely lost, due to changing sea-levels (see *Figure 14*).

Shallow genetic structure detected in almost all sub-Antarctic and Antarctic penguin taxa

Previous studies have suggested that the APF and the STF are barriers to dispersal within penguin populations (Friesen, 2005; Friesen *et al*., 2007; Munro & Burg, 2017; Clucas *et al*., 2018; Frugone *et al*., 2018), which may have led to pronounced genetic structure, particularly within *E. filholi*, *E. chrysolophus chrysolophus*/*E. c. schlegeli*, *Pygoscelis papua*, *Aptenodytes forsteri* and *A. patagonicus*. While some relatively shallow genetic differentiation has been detected among some Southern Ocean populations of these taxa (see *Chapter 3*), these structures have probably only evolved recently (e.g. post-glacially), and the extent of differentiation among them is unlikely to be consistent with long-term isolation, as discussed below.

New data from this thesis revealed no evidence to suggest any population genetic structure in the endemic and localised *E. pachyrhynchus* and *E. robustus* populations. Similarly, there is little evidence to support population level genetic structure between *Eudyptes chrysolophus chrysolophus* and *E. c. schlegeli*. Very shallow genetic structure within these subspecies (when they were combined as a single taxon) appears to correspond to (1) *E. c. schlegeli* north of the APF (Macquarie and Marion Island), (2) *E. c. chrysolophus* north and on the APF (Marion, Prince Edward, Kerguelen and Crozet Islands) and (3) *E. c. chrysolophus* south of the APF (Elephant, South Georgia and South Sandwich Islands). However, this pattern may also be explained by isolation by distance (Frugone *et al*., 2019) and island emergence (see above). Very subtle genetic structure may also be present between the NZ sub-Antarctic population of *E. filholi*, compared to other sub-Antarctic populations. There are potentially slightly higher

levels of population genetic structure present in *E. chrysocome*, however this thesis only sampled two geographically proximate populations on the Falkland Islands. Likewise, population genetic structure may be present between *E. moseleyi* populations on Gough and Amsterdam Island (as suggested by Frugone *et al*., 2018), however this thesis only sampled eight *E. moseleyi* individuals. These shallow patterns within *E. filholi*, *E. chrysocome* and *E. moseleyi* may instead relate to recent isolation during glaciation. Clucas *et al*. (2018) demonstrated negligible genetic structure in *Pygoscelis antarctica* and *P. adeliae*. Similarly, Cristofari *et al*. (2016) demonstrated that *A. forteri* represents a single panmictic unit, despite Younger *et al*. (2017) and Clucas *et al*. (2018) arguing for four distinct metapopulations of *A. forsteri*, and Clucas *et al*. (2016) and Clucas *et al*. (2018) arguing for two distinct metapopulations of *A. patagonicus*. While very shallow genetic structure within *Pygoscelis antarctica*, *P. adeliae*, *Aptenodytes forsteri* and *A. patagonicus* is possible, there remains little evidence for major within-species population genetic structure in these species.

By contrast, populations of *Pygoscelis papua* within the Scotia Arc are distinct (Clucas *et al*., 2014; Clucas *et al*., 2018). This pattern is potentially explained by both population isolation on both sides of the APF, as well as isolation during recent glaciation. An alternative explanation, as put forward by Clucas *et al*. (2018) relates to the behaviour and ecology of *P. papua*. Individuals of *P. papua* forage inshore on locally available prey, rarely ranging beyond the continental shelf (Lescroël & Bost, 2005; Ratcliffe & Trathan, 2012). The tendency of *P. papua* to remain in shelf waters may contribute to the high genetic differentiation by reducing introgression with other populations. Similarly, the coastal lifestyle observed in *Megadyptes antipodes antipodes* may also explain the genetic structure observed within the subspecies, and may explain the population differentiation that has also been observed within the extinct *M. a. waitaha* (Boessenkool *et al*., 2008) and the extant *M. a. antipodes* (Boessenkool *et al*., 2009).

While *Pygoscelis*, *Aptenodytes*, *Megadyptes* and *Eudyptes* penguins have high dispersal capability, and commonly cross the APF or the STF, it remains unlikely that these ocean currents represent significant barriers to dispersal in any of these taxa.

Penguins recolonised sub-Antarctic and Antarctic islands following the last ice age

This thesis provides the first multi-species analysis in Southern Ocean assemblages, demonstrating that rapid climate change promoted recolonisation of sub-Antarctic and Antarctic coastlines. This thesis demonstrates concerted population expansion events in almost all penguin taxa that breed on islands that were impacted by LGM winter sea-ice. In contrast, *E. pachyrhynchus*, *E. robustus*, *E. moseleyi* and *E. chrysocome*, which breed on islands north

of the LGM winter sea-ice, did not experience population expansions. These findings may support Fraser *et al*. (2011), who suggested that islands and major land masses located north of the LGM winter sea-ice, such as NZ, Campbell Island, Antipodes Islands, Auckland, Islands Amsterdam Island, Gough Island, Falkland Islands and Southern South America were likely refugial regions during the LGM (Fraser *et al*., 2009; Fraser *et al*., 2011; Rains *et al*., 2011; Hodgson *et al*., 2014; Fraser *et al*., 2018; Rainsley *et al*., 2019). However, sea levels have risen 120 – 135 m since the LGM (Clark & Mix, 2002), so sub-fossil evidence from these coastlines is likely lost.

Stable climate history promotes high species richness and phylogeographic structure in NZ penguins

Evolutionary processes such as genetic drift are pronounced in isolated island populations. However, glaciation during ice ages can erode such processes, resulting in extinction, biogeographic shifts and post-glacial introgression. While the now-panmictic sub-Antarctic penguin populations may have once been geographically isolated from one another, many penguin populations were impacted by extensive glaciation during the last ice age. As such, these populations likely expanded to mid latitude refugia during past ice ages, and during post-glacial periods recolonised the high latitudes. However, the NZ region represents hotspots for biodiversity both within and among coastal vertebrate lineages. The NZ region encompasses six island archipelagos (NZ South and North Islands, The Snares and Western Chain, Antipodes and Bounty Islands, Auckland Island, Campbell Island, Chatham Islands), which all remained free from substantial glaciation during the last ice age. For this reason, NZ represents a unique geographic region that is still shaped by long-term evolutionary processes.

Almost half of all modern penguins inhabit the NZ region, with pronounced species richness and island-endemism, and distinct phylogeographic patterns within lineages. Four *Eudyptes* species are endemic breeders to NZ islands, separated by distinct biogeographic regions, e.g. *E. pachyrhynchus* on the NZ South Island, *E. robustus* on The Snares and Western Chain, *E. sclateri* on the Antipodes and Bounty Islands and *E. warhami* on the Chatham Islands, with shallow phylogeographic structure separating the Auckland, Antipodes and Campbell Island *E. filholi* lineage from the other non-NZ *E. filholi* lineage. Within *Megadyptes* penguins, there are also distinct biogeographic patterns within regions, e.g. *M. a. antipodes* on Antipodes and Auckland Islands, *M. a. waitaha* on the NZ South and North Island, and *M. a. richdalei* on the Chatham Islands. Even within *Megadyptes* subspecies, there is distinct phylogeographic structure, e.g. three distinct *M. a. antipodes* lineages, two distinct *M. a. waitaha* lineages. Even at a morphological scale, there appears to be distinct, recently evolved phenotypes, e.g. white-

flippered *Eudyptula minor* on the NZ Banks Peninsula, compared to the other *E. minor* phenotype in the rest of NZ and Chatham Islands.

The unique conditions in the NZ region, compared to other Southern Ocean regions may potentially provide some explanation to the high endemism, species richness, phylogeographic structure and morphological diversity. These factors likely include: (1) the geographical location of these islands, which were free from LGM sea-ice; (2) the relatively temperate climate, ranging from the sub-Antarctic Campbell Island to the oceanic-temperate NZ North Island, which may provide optimal climates for penguin populations; and (3) the diverse geological history of islands, e.g. the young Antipodes Islands, Chatham Islands and The Snares compared to the older NZ, Campbell and Auckland Islands, which together may have facilitated founder speciation and subsequent diversification. These abiotic processes may have led to the substantial primary productivity within this region, and may have initiated the high philopatry and localised dispersal within these penguin lineages, as has been noted in *Megadyptes* penguins.

Demography and extinction in temperate penguins

Human settlement initiated the extinction of penguins

Several past studies (Boessenkool *et al*., 2008; Grosser *et al*., 2015; Rawlence *et al*., 2015a; Grosser *et al*., 2016; Rawlence *et al*., 2018) have analysed aDNA sequences from sub-fossil bones to investigate the direct impacts of Polynesian arrival on NZ penguin populations. Boessenkool *et al*. (2008) uncovered the now extinct *M. antipodes waitaha* in the NZ South Island, and Rawlence *et al*. (2018) argued that this subspecies also inhabited the NZ North Island. Similarly, Grosser *et al*. (2016) revealed the local extinction of *Eudyptula minor* on the NZ Otago coastline. Based on radiocarbon dating, the extinction of both *Megadyptes antipodes waitaha* and the Otago population of *Eudyptula minor* appear to correlate with the period immediately following initial human settlement, indicating that human hunting may have played a key role. This thesis builds upon those previous studies by examining possible extinction and lineage reduction events in *Eudyptes* and additional *Megadyptes* taxa. Results from this thesis uncovered two extinct penguins from the Chatham Islands, *Eudyptes warhami* and *Megadyptes antipodes richdalei*. While *Eudyptes warhami* and *Megadyptes antipodes richdalei* were likely hunted to extinction following human arrival to the Chatham Islands (based on associated radiocarbon dates and presence in midden deposits), Travers and Travers (1872) apparently held a live *Eudyptes* penguin from the Chatham Islands in captivity for several days. It is possible, however, that this captive penguin was a vagrant of another

Eudyptes species. While the persistence of *E. warhami* to European times is therefore unlikely, this possibility cannot yet be rejected.

Clearly, the widespread extinction of *E. warhami*, *Megadyptes antipodes waitaha*, *M. a. richdalei* and the local extinction of *Eudyptula minor* in the NZ region alone, during the last few hundred years, demonstrates the vulnerability of isolated Southern Ocean islands to biodiversity loss (see also Cole & Wood, 2017b; Seersholm *et al*., 2018). While studies based on single observations have proposed additional Holocene extinct penguin taxa e.g. Larsen's penguin (*Eudyptes* sp.) and Commerson's Penguin (*Aptenodytes torquata*), Hume (2017) strongly suggests that these suggested taxa are invalid (as shown in this thesis for *Tasidyptes hunteri*).

The Chatham Islands were not re-colonised following extinction events

Studies of several NZ coastal vertebrate taxa have revealed evidence for rapid extinction and recolonisation events, e.g. lineage turnover in *Phocarctos* sea-lions (Collins *et al*., 2014a), *Megadyptes* and *Eudyptula* penguins on the NZ South Island (Boessenkool *et al*., 2008; Rawlence *et al*., 2015a; Grosser *et al*., 2016), yet data from this thesis did not detect any prehistoric range-shift in *Eudyptes* penguins on the NZ mainland or the Chatham Islands. This finding is surprising, as the widespread extinction of *E. warhami* and *Megadyptes antipodes richdalei* on the Chatham Islands would have left niches available for dispersing and founding *Eudyptes sclateri* and *Megadyptes antipodes antipodes* individuals. Both *Eudyptes sclateri* and *Megadyptes antipodes antipodes* are commonly observed as vagrants to the Chatham Islands and NZ shores (Miskelly *et al*., 2006; Robertson *et al*., 2017).

One possible explanation for the lack of Chatham Islands recolonization may be linked to geography. The relatively flat Chatham Islands coastline, which is easy to access, was rapidly and substantially modified by early Polynesian and later European settlers (Towns & Ballantine, 1993; Tennyson & Millener, 1994; Roberts *et al*., 2007). Landscape modification may have blocked the establishment of dispersing founder individuals to the Chatham Islands.

Habitat preferences buffered species from extinction

There is overwhelming evidence demonstrating that the collective impacts of Polynesian and European settlement in NZ, such as hunting, habitat modification and the introduction of exotic predators, have resulted in widespread and local extinctions (Worthy & Holdaway, 2002), and population reductions (Seersholm *et al*., 2018). While previous studies suggested the possibility of range reductions in *Eudyptes pachyrhynchus*, both for the upper NZ South Island and the

lower NZ North Island (Worthy, 1997; Rawlence *et al*., 2018), results from this thesis did not detect any evidence of a range reduction event in the species.

The extinction of *E. warhami, Megadyptes antipodes richdalei*, *M. a. waitaha* and the Otago *Eudyptula minor* lineages occured predominantly on flat, easily accessible landscapes, e.g. Chatham Islands and the NZ east coast. However, the habitat preference of *Eudyptes pachyrhynchus*, which breeds in the steep and more difficult to access forests of NZ's west coast, probably helped buffer the species from the impacts of human arrival.

Recommendations

Phylogenetics and systematics of modern penguins

There are several phylogenetic issues that have yet to be conclusively resolved for modern penguin lineages. While results from this thesis using mitogenomic data suggests that *Aptenodytes* and *Pygoscelis* are sister groups, and basal to all other modern penguins, a recent study using high-coverage nuclear and mitogenomes by Pan *et al*. (in press) demonstrates that *Aptenodytes* alone is basal to all other modern penguins, providing intriguing evidence for the adaptation and biogeographic patterns of Antarctic-evolved penguin genera (*Aptenodytes* and *Pygoscelis*). The relationships of some lineages in this thesis also remain weakly supported. Specifically, low posterior probabilities were observed for the placement of *Eudyptes chrysolophus chrysolophus*/*E. c. schlegeli* within the *Eudyptes* clade, and within *Megadyptes*. While the recent penguin phylogeny by Pan *et al*. (2019) did not include any extinct penguin taxa, it does support the mitogenome topology presented in this thesis (with the exception of *Aptenodytes*/*Pygoscelis* relationships). Specifically, Pan *et al*. (2019) demonstrates that *Eudyptes chrysolophus chrysolophus* and *E. c. schlegeli* are basal and sister to all other *Eudyptes* penguins, despite the weakly supported branch presented in this thesis.

In addition, the stem penguin *Waimanu manneringi* could not be used as a calibration point, and *Phoebastria albatrus* could not be used as the outgroup, because the effective priors for the crown nodes and the split between Sphenisciformes and Procellariformes became unrealistic when they were used. This could be due to a rate slowdown in the large-bodied penguin clade (Ksepka & Phillips, 2015) and may be exacerbated by the long branch separating crown penguins from Procellariiformes (see *Chapter 2*). These relationships may be further addressed under a future calibrated phylogenetic framework using phylogenomic methods in Zhang, (2015), a newly developed full genome dataset of penguins in Pan *et al*. (in press) and improved ways to combine molecular and morphological characters (e.g. Gavryushkina *et al*., 2017).

Similarly, the systematics of some penguin taxa remain unclear, and require further addressing under a phylogenetic framework that incorporates whole genomes and morphology. For example, it remains unlikely that *Eudyptes chrysolophus chrysolophus* and *E. c. schlegeli* represent separate species. Similarly, the very localised *Eudyptula minor albosignata* on NZ's Banks Peninsula is not genetically distinctive and seems unlikely to warrant recognition as either a separate species or subspecies. In both cases, the local 'taxa' have been based on single plumage characters, without any additional convincing corroborating evidence. The relationships between the three *Megadyptes* subspecies also remain unclear. Although size can distinguish the three subspecies, phylogenetic analyses often generated low posterior probabilities, suggesting that the three clades may have diverged almost simultaneously. While other widespread penguins apparently lack comparable phylogeographic differentiation, *Pygoscelis papua*, for example, shows pronounced population differentiation within the Scotia Arc. Clucas *et al*. (2014) and Clucas *et al*. (2018) suggested these lineages could be recognised as distinct subspecies. In the absence of congruent morphological distinctions, however, it may be more appropriate to recognise *P. papua* populations under an Evolutionary Significant Unit or Management Unit framework (as was put forward by Rexer-Huber *et al*., in press for *Procellaria aequinoctialis*). Furthermore, it remains possible that additional population differentiation exists among Atlantic (Bouvet), Indian (Prince Edward Islands/Marion Island, Crozet, Kerguelen) and Pacific (Macquarie Island) *P. papua* colonies. Future population and systematic assessments based on whole genomes, behaviour and ecology will be critical for taxonomy, and will ensure future conservation and management planning of sub-Antarctic and Antarctic penguin populations.

Evolution and biogeography of modern penguins

Concerted population expansions and biogeographical shifts following climate events

This thesis examined genome-wide population structure in all but one *Eudyptes* species (*E. sclateri*), and expanded on past studies of *Aptenodytes* and *Pygoscelis*, by inferring population expansion events following the LGM. While previous studies have explored population genetic structure to some detail within *Eudyptula*, *Megadyptes*, *Spheniscus* and *Eudyptes sclateri*, no study has undertaken a population genomic approach in studying these taxa. Future analyses may include SNPs or whole genomes for those understudied taxa.

Future demographic analyses may also benefit from interrogating SNPs or whole genomes (Pan *et al*., in press) across all penguin taxa, including unsampled northern taxa, such as *E. sclateri*, *E. warhami*, *Megadyptes antipodes antipodes*, *M. a. waitaha*, *M. a. richdalei*, *Eudyptula minor*,

E. novaehollandiae, *Spheniscus demersus*, *S. magellanicus*, *S. mendiculus* and *S. humboldti*. Such an analysis may test whether the patterns revealed in this thesis are consistent among other northerly-distributed taxa across the entire Southern Ocean.

Predicting future biological responses to climate change

Results from this thesis suggest that many sub-Antarctic and Antarctic penguin taxa may have expanded south from temperate refugia following the LGM, with population sizes expanding as global climate warmed. However, penguins are sensitive to climate change, and current global warming is resulting in declining populations of some taxa (e.g. Barbraud & Weimerskirch, 2001; Forcada *et al*., 2006).

Some lineages within a species may be better adapted to warmer environments than others. For example, *Megadyptes antipodes antipodes* probably diverged from *M. a. waitaha* or *M. a. richdalei* on the cold sub-Antarctic islands, and as this lineage has recently expanded north to mainland NZ, it may not be adapted to the warmer NZ climate (Rawlence *et al*., 2019). Similarly, the widespread *Eudyptes chrysolophus chrysolophus*/*E. c. schlegeli*, *Aptenodytes patagonicus* and *Pygoscelis papua* all breed on islands both south and north of the APF. While most of these taxa exhibit little or no population differentiation, populations north of the APF may be better adapted to warmer environments than those south of the APF. Northern populations, therefore, could have greater potential to adapt to future climate change. New population-wide data across whole genomes may provide such clues into thermal adaptation.

Conservation of penguin populations

Many penguin populations are already declining, or are predicted to decline in the near future (Frost *et al*., 1976; Hays, 1984; Randall & Randall, 1986; Dann, 1991; Cunningham & Moors, 1994; Dann *et al*., 2000; Kemper *et al*., 2001; Forcada *et al*., 2006; Sander *et al*., 2007; Le Bohec *et al*., 2008; Cuthbert *et al*., 2009; Jenouvrier *et al*., 2009; Robson *et al*., 2011; Trivelpiece *et al*., 2011; King *et al*., 2012; Lynch *et al*., 2012; Sherley *et al*., 2013; Hiscock & Chilvers, 2014; Jenouvrier *et al*., 2014; Mattern *et al*., 2017; Mattern & Wilson, 2018; Ropert-Coudert *et al*., 2019; Boersma *et al*., in press). For example, *Eudyptes filholi* on Campbell Island have declined by more than 99.4% between 1942 – 1985 (Cunningham & Moors, 1994), *Pygoscelis adeliae* on Penguin Island have declined by 75% between 1979 – 2004 (Sander *et al*., 2007) and *Pygoscelis antarctica* on Penguin Island have declined by 66% between 1979 – 2004 (Sander *et al*., 2007), while *Aptenodytes forsteri* are predicted to decline by 19% by 2100 (Jenouvrier *et al*., 2014) and *Megadyptes antipodes antipodes* are predicted to be extinct on the NZ mainland by 2043 (Mattern *et al*., 2017). These losses have been linked to disease, predation

by introduced mammals, competition with fisheries and mortality associated with fishery bycatch, pollution, declines in key prey species, reductions in sea-ice, warming sea-surface temperatures and global climate change (Trivelpiece *et al*., 2011; Lynch *et al*., 2012; Mattern *et al*., 2017; Herrah *et al*., 2019).

Current conservation actions for many penguin populations on remote Southern Ocean islands are limited to occasional population counts (Mattern & Wilson, 2018). However, in addition to ongoing monitoring of penguin populations, it is also necessary to understand population genetic connectivity and demographic histories (crucial for conservation management planning), reduce the impacts of introduced predatory mammals and fisheries, restore ecosystems and mitigate climate change (Forcada *et al*., 2006; Forcada & Trathan, 2009; Keogan *et al*., 2018; Mattern & Wilson, 2018; Heerah *et al*., 2019). While these threats pose a grave risk to most penguin populations, understanding penguin evolution and biogeography, as was done in this thesis, provides baseline data that can help guide conservation efforts.

Conclusions

This thesis aimed to test whether the classic patterns of island biology observed on local archipelagos could be scaled up to the islands in the vast Southern Ocean. Focussing on penguins, an order of seabirds that rely on Southern Ocean islands to breed, this thesis examined patterns of evolution, biogeography and extinction, across three temporal scales (millions of years, thousands of years and hundreds of years).

The main results from this thesis were (1) the emergence of geologically-young islands in the Southern Ocean played a key role in the diversification of island-endemic penguin species, including *Spheniscus mendiculus*, *Eudyptes moseleyi*, *E. chrysolophus schlegeli*, *E. warhami* and possibly *E. robustus* and *Megadyptes antipodes richdalei*; (2) post-LGM climate warming led to rapid recolonisation and reduced genetic diversity across all widespread penguin species in the sub-Antarctic and Antarctic region, including *Eudyptes filholi*, *E. chrysolophus chrysolophus*/*E. c. schlegeli*, *Aptenodytes patagonicus*, *Pygoscelis papua*, *P. antarctica* and *P. adeliae* while the ice-adapted *A. forsteri* began expanding slightly earlier, during the LGM; (3) recent Polynesian arrival to NZ and the Chatham Islands led to the local and widespread extinction of penguins, including the extinction of two newly described taxa, the large *Eudyptes warhami* and the dwarf *Megadyptes antipodes richdalei*, along with *M. a. waitaha* and a local population of *Eudyptula minor*; and (4) geography of Southern Ocean islands, combined with unique habitat preferences buffered some penguin species from local extinction following climate and anthropogenic change, including *Eudyptes chrysocome* and *E. pachyrhynchus*.

Results from this thesis suggest that classic processes underpinning island biogeography and evolution in archipelagos worldwide, such as island emergence (*Chapter 2*) and founder speciation (*Chapter 2*, *Chapter 3*) have played important roles in diversification of isolated Southern Ocean biota. Results also demonstrate that Southern Ocean island biota are particularly sensitive to both climate change and anthropogenic impacts, which together have led to widespread extinctions (*Chapter 2*, *Chapter 4*), refugia (*Chapter 3*, *Chapter 4*) and range expansions (*Chapter 3*). Taken together, this thesis provides the broad genomic evidence for rapid evolutionary and biogeographic shifts in the iconic penguin fauna of the vast Southern Ocean.

References

Acosta Hospitaleche C, Tambussi C, Donato M & Cozzuol M. (2007). A new Miocene penguin from Patagonia and its phylogenetic relationships. *Acta Palaeontologica Polonica*, 52(2): 299 – 314.

Adamson DA, Selkirk PM, Price DM & Selkirk JM. (1996). Pleistocene uplift and palaeoenvironments of Macquarie Island: evidence from palaeobeaches and sedimentary deposits. In *Papers and Proceedings of the royal society of Tasmania*, 130(2): pp. 25 – 32.

Akaike H. (1974). A new look at the statistical model identification. In *Selected Papers of Hirotugu Akaike.* Springer, New York, NY.

Allendorf FW. (2016). Genetics and the conservation of natural populations: allozymes to genomes. *Molecular Ecology*, 26(2): 420 – 430.

Amadon D. (1949). The seventy-five percent rule for subspecies. *The Condor*, 51(6): 250 – 258.

Anderson AJ. (2017). Changing perspectives upon Māori colonisation voyaging. *Journal of the Royal Society of New Zealand*, 47(3): 222 – 231.

Andrews KR, Good JM, Miller MR, Luikart G & Hohenlohe PA. (2016). Harnessing the power of RADseq for ecological and evolutionary genomics. *Nature Reviews Genetics,* 17(2): 81 – 92.

Angel A, Wanless RM & Cooper J. (2008). Review of impacts of the introduced house mouse on islands in the Southern Ocean: are mice equivalent to rats? *Biological Invasions*, 11(7): 1743-1754.

Armour KC, Marshall J. Scott JR, Donohoe A. & Newsom ER. (2016). Southern Ocean warming delayed by circumpolar upwelling and equatorward transport. *Nature Geoscience*, 9(7): 549.

Armstrong PH. (1994). Human impact on the Falkland Islands environment. *Environmentalist*, 14(3): 215 – 231.

Bacon CD, Baker WJ & Simmons MP. (2012). Miocene dispersal drives island radiations in the palm tribe Trachycarpeae (Arecaceae). *Systematic Biology*, 61(3): 426 – 442.

Baird NA, Etter PD, Atwood TS, Currey MC, Shiver AL, Lewis ZA, Selker EU, Cresko WA & Johnson EA. (2008). Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PloS One*, 3(10): e3376.

Baker A & Marshall H. (1997). Mitochondrial control region sequences as tools for understanding evolution. In: Avian molecular evolution and systematics. Academic Press, New York.

Baker AJ, Pereira SL, Haddrath OP & Edge KA. (2006). Multiple gene evidence for expansion of extant penguins out of Antarctica due to global cooling. *Proceedings of the Royal Society London B Biological Series*, 273(1582): 11 – 17.

Bandelt H, Forster P & Röhl A. (1999). Median-joining networks for inferring intraspecific phylogenies. *Molecular Biolology and Evolution*, 16(1): 37 – 48.

Banks J, Van Buren A, Cherel Y & Whitfield JB. (2006). Genetic evidence for three species of rockhopper penguins, *Eudyptes chrycosome*. *Polar Biology*, 30(1): 61 – 67.

Barbraud C & Weimerskirch H. (2001). Emperor penguins and climate change. *Nature*, 411(6834): 183 – 186.

Barker PF. & Burrell J. (1977). The opening of Drake passage. *Marine Geology*, 25(1 – 3): 15 – 34.

Beaumont MA, Zhang WY & Balding DJ. (2002). Approximate Bayesian computation in population genetics. *Genetics*, 162(4): 2025 – 2035.

Benavides E, Baum R, Snell HM, Snell HL & Sites JW. (2009) Island biogeography of Galápagos lava lizards (Tropiduridae: Microlophus): species diversity and colonization of the archipelago. *Evolution: International Journal of Organic Evolution*, 63(6): 1606 – 1626.

Bergner LM, Dussex N, Jamieson IG & Robertson BC. (2016). European colonization, not Polynesian arrival, impacted population size and genetic diversity in the critically endangered New Zealand kākāpō. *Journal of Heredity*, 107(7): 593 – 602.

Bergstrom DM & Chown SL. (1999). Life at the front: history, ecology and change on Southern Ocean Islands. *Trends in Ecology and Evolution.* 14(12): 472 – 477.

Bertelli S & Giannini NP. (2005). A phylogeny of extant penguins (Aves: Sphenisciformes) combining morphology and mitochondrial sequences. *Cladistics*, 21(3): 209 – 239.

BirdLife International. (2016). *Eudyptes pachyrhynchus*. IUCN Red List of Threatened Species. doi: http:/dx.doi.org/10.2305/IUCN.UK.2016-3.RLTS. T22697776A93638571.en.

Boast AP, Chapman B, Herrera MB, Worthy TH, Scofield RP, Tennyson AJD, Houde P, Bunce M, Cooper A & Mitchell KJ. (2019). Mitochondrial genomes from New Zealand's extinct adzebills (Aves: Aptornithidae: *Aptornis*) support a sister-taxon relationship with the Afro-Madagascan Sarothruridae. *Diversity*, 11(2): 24.

Boersma PD. (2008). Penguins as marine sentinels. *BioScience*, 58(7): 597 – 607.

Boersma PD, García Borboroglu P, Gownaris NJ, Bost CA, Chiaradia A, Ellis S, Schneider T, Seddon PJ, Simeone A, Trathan PN, Waller LJ & Wienecke B. (in press). Applying science to pressing conservation needs for penguins. *Conservation Biology.*

Boessenkool S, Austin JA, Worthy TH, Scofield P, Cooper A, Seddon PJ & Waters JM. (2008). Relict or colonizer? Extinction and range expansion of penguins in southern New Zealand. *Proceedings of the Royal Society B Biological Series*, 276(1658): 815 – 821.

Boessenkool S, Star B, Waters JM & Seddon PJ. (2009). Multilocus assignment analyses reveal multiple units and rare migration events in the recently expanded yellow-eyed penguin (*Megadyptes antipodes*). *Molecular Ecology*, 18(11): 2390 – 2400.

Boessenkool S, Star B, Seddon PJ & Waters JM. (2010). Temporal genetic samples indicate small effective population size of the endangered yellow-eyed penguin. *Conservation genetics*, 11(2): 539 – 546.

Bollmer JL, Kimball RT, Whiteman NK, Sarasola JH & Parker PG. (2006). Phylogeography of the Galápagos hawk (*Buteo galapagoensis*): a recent arrival to the Galápagos Islands. *Molecular Phylogenetics and Evolution*, 39(1): 237 – 247.

Borboroglu PG & Boersma PD. (2013). Penguins: Natural History and Conservation. University of Washington Press.

Bost CA, Cotté C, Bailleul F, Cherel Y, Charrassin JB, Guinet C, Ainley DG. & Weimerskirch H. (2009). The importance of oceanographic fronts to marine birds and mammals of the southern oceans. *Journal of Marine Systems*, 78(3): 363 – 376.

Bouckaert R & Heled J. (2014). DensiTree2: seeing trees through the forest. *BioRxiv*, 012401.

Bouckaert R, Heled J, Kühnert D, Vaughan T, Wu C-H, Xie D, Suchard MA, Rambaut A & Drummond AJ. (2014). BEAST 2: A Software Platform for Bayesian Evolutionary Analysis. *PLoS Computational Biology*, 10(4): e1003537.

Bouzat JL, Walker BG & Boersma PD. (2009). Regional genetic structure in the Magellanic penguin (*Spheniscus magellanicus*) suggests metapopulation dynamics. *The Auk*, 126(2): 326 – 334.

Brambati A, Melis R, Quaia T & Salvi G. (2002). Late Quaternary climatic changes in the Ross Sea area, Antarctica. In Antarctica at the Close of the Millennium*. Royal Society New Zealand*, Wellington.

Brand L, Urbina M, Chadwick A, DeVries TJ & Esperante R. (2011). A high resolution stratigraphic framework for the remarkable fossil cetacean assemblage of the Miocene/Pliocene Pisco Formation, Peru. *Journal of South American Earth Sciences*, 31(4): 414 – 425.

Bried J. (2003). Impact of vagrant predators on the native fauna: A Short-eared Owl (*Asio flammeus*) preying on Madeiran Storm Petrels (*Oceanodroma castro*) in the Azores. *Life and Marine Sciences*, 20A: 57 – 60.

Bronk Ramsey C. (2017). Methods for Summarizing Radiocarbon Datasets. *Radiocarbon*, 59(2): 1809 – 1833.

Brotherton P, Endicott P, Sanchez JJ, Beaumont M, Barnett R, Austin J & Cooper A. (2007). Novel high-resolution characterisation of ancient DNA reveals C > U-type base modification events as the sole cause of post mortem miscoding lesions. *Nucleic Acids Research*, 35(17): 5717 – 5728.

Brotherton P, Haak W, Templeton J, Brandt G, Soubrier J, Adler CJ, Richards SM, Der Sarkissan C, Ganslmeier R, Friederich S, Dresely V, van Oven M, Kenyon R, Van der Hoek MB, Korlach J, Luong K, Ho SYW, Quintana-Murci L, Behar DM, Alt KW, Cooper A & The Genographic Consortium. (2013). Neolithic mitochondrial haplogroup H genomes and the genetic origins of Europeans. *Nature Communications*, 4: 1764.

Brown CR & Klages NT. (2019). Seasonal and annual variation in diets of Macaroni (*Eudyptes chrysolophus chrysolophus*) and Southern rockhopper (*E. chrysocome chrysocome*) penguins at sub-Antarctic Marion Island. *Journal of Zoology*, 212(1): 7 – 28.

Buckley SJ, Domingos FMCB, Attard CRM, Brauer CJ, Sandoval-Castillo J, Lodge R, Unmack PJ & Beheregaray LB. (2018). Phylogenomic history of enigmatic pygmy perches: implications for biogeography, taxonomy and conservation. *Royal Society open science*, *5*(6): 172125

Burg TM & Croxall JP. (2001) Global relationships amongst Black-browed and Grey-headed Albatrosses: Analysis of population structure using mitochondrial DNA and microsatellites. *Molecular Ecology*, 10(11): 2647 – 2660.

Burger J & Gochfeld M. (2004). Marine birds as sentinels of environmental pollution. *EcoHealth*, 1(3): 263 – 274.

Burridge CP, Peucker AJ, Valautham SK, Styan CA & Dann P. (2015). Nonequilibrium conditions explain spatial variability in genetic structuring of little penguin (*Eudyptula minor*). *Journal of Heredity*, 106(3): 228 – 237.

Caccone A, Gibbs JP, Ketmaier V, Suatoni E & Powell JR. (1999). Origin and evolutionary relationships of giant Galápagos tortoises. *Proceedings of the National Academy of Sciences USA*, 96(23): 13223 – 13228.

Campbell HJ, Begg J, Bea A, Carter B, Curtis N, Davies G, Emberson R, Given D, Goldberg J, Holt K, Hoernli K, Malahoff A, Mildenhall D, Landis C, Paterson A & Trewick S. (2008). Geological consideration relating to the Chatham Islands, mainland New Zealand, and the history of New Zealand terrestrial life. *Geolological Society of New Zealand Miscellaneous Publication*, 126: 126.

Carøe C, Gopalakrishnan S, Vinner L, Mak SST, Sinding MHS, Samaniego JA, Wales N, Sicheritz-Pontén T, Gilbert MTP. (2018). Single-tube library preparation for degraded DNA. *Methods in Ecology and Evolution*, 9(2): 410 – 419.

Carrea C, Burridge CP, Wienecke B, Emmerson LM, White D & Miller KJ. (2019). High vagility facilitates population persistence and expansion prior to the Last Glacial Maximum in an Antarctic top predator: the snow petrel (*Pagodroma nivea*). *Journal of Biogeography*, 46(2): 442 – 453.

Ceballos G, Ehrlich PR, Barnosky AD, García A, Pringle RM & Palmer TM. (2015). Accelerated modern human-induced species losses: entering the sixth mass extinction. *Science Advances*, 1(5): e1400253.

Challies CW & Burleigh RR. (2004). Abundance and breeding distribution of the white-flippered penguin (*Eudyptula minor albosignata*) on Banks Peninsula, New Zealand. *Notornis*, 51(1): 1 – 6.

Chau JH, Born C, McGeoch MA, Bergstrom D, Shaw J, Terauds A, Mairal M, Le Roux JJ, Jansen van Vuuren B. (2019). The influence of landscape, climate, and history on spatial genetic patterns in keystone plants (*Azorella*) on sub-Antarctic islands. *Molecular Ecology*. 28(14): 3291 – 3305.

Chávez Hoffmeister MC. (2014). Phylogenetic characters in the humerus and tarsometatarsus of penguins. *Polish Polar Research*, 35(3): 469 – 496.

Chávez Hoffmeister MC, Briceño JDC & Nielsen SN. (2014). The evolution of seabirds in the Humboldt Current: new clues from the Pliocene of central Chile. *Plos One*, 9(3): e90043.

Chen I-C, Hill JK, Ohlemuller R, Roy DB & Thomas CD. (2011). Rapid range shifts of species associated with high levels of climate warming. *Science*, 333(6045): 1024 – 1026.

Chown SL, Gremmen NJM. & Gaston KJ. (1998). Ecological biogeography of southern ocean islands: species-area relationships, human impacts, and conservation. *The American Naturalist*, 152(4): 562 – 575.

Christidis L & Boles WE. (2008). Systematics and Taxonomy of Australian Birds. *CSIRO Publishing*, Canberra.

Clark PU & Mix AC. (2002). Ice sheets and sea level of the Last Glacial Maximum. *Quaternary Science Reviews*, 21(1 – 3): 1 – 7.

Clarke JA, Olivero EB & Puerta P. (2003). Description of the earliest fossil penguin from South America and first Paleogene vertebrate locality of Tierra del Fuego, Argentina. *American Museum Novitates*, 1 – 18.

Clegg SM, Degnan SM, Kikkawa J, Moritz C, Estoup A & Owens IPF. (2002). Genetic consequences of sequential founder events by an island-colonizing bird. *Proceedings of the National Academy of Sciences USA*, 99(12): 8122 – 8132.

Clucas GV, Dunn MJ, Dyke G, Emslie SD, Levy H, Naveen R, Polito MJ, Pybus OG, Rogers AD & Hart T. (2014). A reversal of fortunes: climate change 'winners' and 'losers' in Antarctic Peninsula penguins. *Nature Scientific Reports*, 4: 5024.

Clucas GV, Younger JL, Kao D, Rogers AD, Handley J, Miller GD, Jouventin P, Nolan P, Gharbi K, Miller KJ & Hart T. (2016). Dispersal in the sub-Antarctic: king penguins show remarkably little population genetic differentiation across their range. *BMC Evolutionary Biology*, 16(1): 211.

Clucas GV, Younger JL, Kao D, Emmerson L, Southwell C, Wienecke B, Rogers AD, Bost C-A, Miller GD, Polito MJ, Lelliot P, Handley J, Crofts S, Phillips RA, Dunn MJ, Miller KJ & Hart T. (2018). Comparative population genomics reveals key barriers to dispersal in Southern Ocean penguins. *Molecular Ecology*, 27(23): 4680 – 4697.

Cole TL, Hammer MP, Unmack PJ, Teske P, Adams M & Beheregaray LB. (2016). Range-wide fragmentation in a threatened fish associated with post-European settlement modification in the Murray-Darling Basin, Australia. *Conservation Genetics*, 17(6): 1377 – 1391.

Cole TL & Wood JR. (2017a). Hybridisation in the last remaining individuals of the extinct Fiordland population of Brown Teal (*Anas chlorotis*). *Records of the Canterbury Museum*, 31: 159 – 165.

Cole TL & Wood JR. (2017b). The ancient DNA revolution: the latest era in unearthing New Zealand's faunal history. *New Zealand Journal of Zoology,* 45(2): 91 – 120.

Cole TL, Waters J, Shepherd LD, Rawlence NJ, Joseph L & Wood JR. (2018). Ancient DNA reveals that the 'extinct' Hunter Island penguin (*Tasidyptes hunteri*) is not a distinct taxon. *Zoological Journal of the Linnean Society*, 182(2): 459 – 464.

Cole TL, Rawlence NJ, Dussex N, Ellenberg U, Houston DM, Mattern T, Miskelly CM, Morrison KW, Scofield RP, Tennyson AJD, Thompson DR, Wood JR & Waters JM. (2019a). Ancient DNA of crested penguins: Testing for temporal genetic shifts in the world's most diverse penguin clade. *Molecular Phylogenetics and Evolution.* 131: 72 – 79.

Cole TL, Ksepka DT, Mitchell KJ, Tennyson AJD, Thomas DB, Pan H, Zhang G, Rawlence NJ, Wood JR, Bover P, Bouzat JL, Cooper A, Fiddaman SR, Hart T, Miller G, Ryan PG, Shepherd LD, Wilmshurst JM & Waters JM. (2019b). Mitogenomes uncover extinct penguin taxa and reveal island formation as a key driver of speciation. *Molecular Biology and Evolution*, 36(4): 784 – 797.

Collins CJ, Rawlence NJ, Prost S, Anderson CNK, Knapp M, Scofield RP, Robertson BC, Smith I, Matisoo-Smith EA, Chilvers BL & Waters JM. (2014a). Extinction and recolonization of coastal megafauna

following human arrival in New Zealand. *Proceedings of the Royal Society B: Biological Sciences*, 281(1786): 20140097.

Collins CJ, Rawlence NJ, Worthy TH, Scofield RP, Tennyson AJD, Smith I, Knapp M & Waters JM. (2014b). Pre-human New Zealand sea lion (*Phocarctos hookeri*) rookeries on mainland New Zealand. *Journal of the Royal Society of New Zealand*, 44(1): 1 – 16.

Convey P & Lebouvier M. (2009). Environmental change and human impacts on terrestrial ecosystems of the sub-Antarctic islands between their discovery and the mid-twentieth century. *Papers and Proceedings of the Royal Society of Tasmania*, 143(1): 33 – 44.

Cooper A & Poinar HN. (2000). Ancient DNA: do it right or not at all. *Science*, 289(5482): 1139.

Cooper RA. (2004). The New Zealand geological timescale, Institute of Geological and Nuclear Sciences Monograph, 22.

Cornuet J-M, Ravigné V & Estoup A. (2010). Inference on population history and model checking using DNA sequence and microsatellite data with the software DIYABC (v1.0). *BMC Bioinformatics*, 11(1): 401.

Cornuet JM, Pudlo P, Veyssier J, Dehne-Garcia A, Gautier M, Leblois R, Marin JM & Estoup A. (2014). DIYABC v2.0: A software to make approximate Bayesian computation inferences about population history using single nucleotide polymorphism, DNA sequence and microsatellite data. *Bioinformatics*, 30(8): 1187 – 1189.

Corrigan LJ, Fabiani A, Chauke LF, McMahon CR, de Bruyn M, Bester MN, Bastos A, Campagna C, Muelbert MMC & Hoelzel AR. (2016). Population differentiation in the context of Holocene climate change for a migratory marine species, the southern elephant seal. *Journal of Evolutionary Biology*, 29(9): 1667 – 1679.

Courtois B, Audebert A, Dardou A, Roques S, Ghneim-Herrera T, Droc G, Frouin J, Rouan L, Gozé E, Kilian A, Ahmadi N & Dingkuhn M. (2013). Genome-wide association mapping of root traits in a Japonica rice panel. *PloS One*, 8(11): e78037.

Cowie RH & Holland BS. (2006). Dispersal is fundamental to biogeography and the evolution of biodiversity on oceanic islands. *Journal of Biogeography*, 33(2): 193 – 198.

Crawford RJM, Whittington PA, Upfold L, Ryan PG, Petersen SL, Dyer BM & Cooper J. (2009). Recent trends in numbers of four species of penguins at the Prince Edward Islands. *African Journal of Marine Science*, 31(3): 419 – 426.

Crawford RJM, Cooper J, Dyer BM, Greyling MD, Klages MTW, Nel DC, Nel JL, Petersen SL & Wolfaardt AC. (2010). Decrease in numbers of the eastern rockhopper penguin *Eudyptes chrysocome filholi* at Marion Island, 1994/95-2002/03. *African Journal of Marine Science*, 25(1): 487 – 498.

Cristofari R, Bertorelle G, Ancel A, Benazzo A, Le Maho Y, Pondanis PJ, Stenseth NC, Trathan PN, Whittington JD, Zanetti E, Zitterbart DP, Le Bohec C & Trucchi E. (2016). Full circumpolar migration ensures evolutionary utility in the Emperor penguin. *Nature Communications*, 7: 11842.

Cristofari R, Liu X, Bonadonna F, Cherel Y, Pistorius P, Le Maho YL, Raybaud V, Stenseth NC, Le Bohec C & Trucchi E. (2018). Climate-driven range shifts of the king penguin in a fragmented ecosystem. *Nature Climate Change*, 8(3): 245.

Cruz VM, Kilian A & Dierig DA. (2013). Development of DArT marker platforms and genetic diversity assessment of the U.S. collection of the new oilseed crop *lesquerella* and related species. *PloS One*, 8(5): e64062.

Csilléry K, François O & Blum MG. (2012). abc: an R package for approximate Bayesian computation (ABC). *Methods in Ecology and Evolution*, 3(3): 475 – 9.

Cunningham DM & Moors PJ. The decline of rockhopper penguins *Eudyptes chrysocome* at Campbell Island, Southern ocean and the influence of rising sea temperatures. *Emu,* 94(1): 27 – 36.

Cuthbert R, Cooper J, Burle MH, Glass CJ, Glass JP, Glass S, Glass T, Hilton GM, Sommer ES, Wanless RM & Ryan PG. (2009). Population trends and conservation status of the Northern Rockhopper Penguin *Eudyptes moseleyi* at Tristan da Cunha and Gough Island. *Bird Conservation International*, 19(1): 109 – 120.

Dabney J, Knapp M, Glocke I, Gansauge M, Weihmann A, Nickel B, Valdiosera C, Garcia N, Pääbo, S, Arsuaga J-L & Meyer M. (2013). Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. *Proceedings of the National Academy of Sciences USA*, 110(39): 15758 – 15763.

DaCosta JM & Sorenson MD. (2016). ddRAD-seq phylogenetics based on nucleotide, indel, and presence-absence polymorphisms: analyses of two avian genera with contrasting histories. *Molecular Phylogenetics and Evolution*, 94: 122 – 135.

Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, Handsaker R, Lunter G, Marth G, Sherry ST, McVean G, Durbin R & 1000 Genomes Project Analysis Group. (2011). The variant call format and VCFtools. *Bioinformatics*, 27(15): 2156 – 2158.

Dann P. (1991). Distribution, population trends and factors influencing the population size of little penguins *Eudyptula minor* on Phillip Island, Victoria. *Emu*, 91(5): 263 – 272.

Dann P, Norman FI, Cullen JM, Neira FJ & Chiaradia A. (2000). Mortality and breeding failure of little penguins, *Eudyptula minor*, in Victoria, 1995 – 96, following a widespread mortality of pilchard, *Sardinops sagax*. *Marine and Freshwater Research*, 51(4): 355 – 362.

Dantas GPM, Maria GC, Marasco ACM, Castro LT, Almeida VS, Santos FR, Oliveira LR, Crespo E, Frere E, Milliones A, González-Acuña D, Morgante JS & Vianna JA. (2018). Demographic history of the Magellanic Penguin (*Spheniscus magellanicus*) on the Pacific and Atlantic coasts of South America. *Journal of Ornithology*, 159(3), 643 – 655.

Dantas GPM, Oliveira LR, Santos AM, Flores MD, de Melo DR, Simeone A, González-Acuña D, Luna-Jorquera G, Le Bohec C, Valdés-Velásquez A, Cardeña M, Morgante JS & Vianna JA. (2019). Uncovering population structure in the Humboldt penguin (*Spheniscus humboldti*) along the Pacific coast at South America. *PLoS One*, 14(5): e0215293.

Darriba D, Taboada GL, Doallo R, & Posada D. (2012). jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods*, 9(8): 772.

Darwin C. (1859). On the origins of species by means of natural selection. *Murray*, London.

Davey JW & Blaxter ML. (2010). RADSeq: next-generation population genetics. *Briefings in Functional Genomics*, 9(5 – 6): 416 – 423.

Davidson J. (1988). The coming of the Maori. In The celebration history of the Kapiti District: 100 years plus. *Kapiti Borough Council*.

Davis LS & Renner M. (2003). Penguins. *T & AD Poyser*, London.

Davis LS. (2013). Erect-crested penguin (*Eudyptes sclateri*). In: Biology and Conservation of the World's penguins. *University of Washington Press*, Seattle.

Davis MB & Shaw RG. (2001). Range shifts and adaptive responses to Quaternary climate change. *Science*, 292(5517): 673 – 679.

de Bruyn M, Hall BL, Chauke LF, Baroni C, Koch PL & Hoelzel AR. (2009). Rapid response of a marine mammal species to Holocene climate and habitat change. *PloS Genetics*, 5(7): e1000554.

de Dinechin M, Ottval R, Quillfeldt P & Jouventin P. (2009). Speciation chronology of rockhopper penguins inferred from molecular, geological and palaeoceanographic data. *Journal of Biogeography*, 36(4): 693 – 702.

de Villiers MS, Cooper J, Carmichael N, Glass JP, Liddle GM, McIvor E, Micol T & Roberts A. (2006). Conservation management at Southern Ocean Islands: towards the development of best-practice guidelines. *Polarforschung*, 75(2 – 3): 113 – 131.

Degrange FJ, Ksepka DT & Tambussi CP. (2018). Redescription of the oldest crown clade penguin: cranial osteology, jaw myology, neuroanatomy, and phylogenetic affinities of *Madrynornis mirandus*. *Journal of Vertebrate Paleontology*, 38(2): e1445636.

de Oliveira EA, Bertollo LAC, Rab P, Ezaz T, Yano CF, Hatanaka T, Jegede OI, Tanomtong A, Liehr T, Sember A, Maruyama SR, Feldberg E, Viana PF & Cioffi MB. (2019). Cytogenetics, genomics and biodiversity of the South American and African Arapaimidae fish family (Teleostei, Osteoglossiformes). *PLoS One*, 14(3): e0214225.

Diamond J. (1990). New Zealand as an archipelago: an international perspective. In: Towns DR, CH Daugherty, IAE Atkinson, editors. Ecological restoration of New Zealand islands. Conservation Sciences Publication No 2. Wellington, New Zealand: Department of Conservation; p. 3 – 8.

Dickinson E & Remsen J. (2013). The Howard & Moore Complete Checklist of the Birds of the World Volume 1. *Aves Press*, Eastbourne.

Drummond AJ, Rambaut A, Shapiro B & Pybus OG. (2005). Bayesian coalescent inference of past population dynamics from molecular sequences. *Molecular Biology and Evolution*, 22(5): 1185 – 1192.

Duchene S, Duchene D, Holmes EC & Ho SYW. (2015). The performance of the date-randomisation test in phylogenetic analyses of time-structured virus data. *Molecular Biology and Evolution*, 32(7): 1895 – 1906.

Duncan RP, Boyer AG & Blackburn TM. (2013). Magnitude and variation of prehistoric bird extinctions in the Pacific. *Proceedings of the National Academy of Sciences USA*, 110(16): 6436 – 6441.

Durbin R, Eddy SR, Krogh A & Mitchison GX. (1998). Biological sequence analysis: probabilistic models of proteins and nucleic acids. *Cambridge university press*, Cambridge.

Dussex N, Rawlence NJ & Robertson BC. (2015). Ancient and contemporary DNA reveal a pre-human decline but no population bottleneck associated with recent human persecution in the kea (*Nestor notabilis*). *PloS ONE*, 10(2): e0118522.

Dussex N, von Seth J, Robertson BC & Dalén L. (2018). Full mitogenomes in the critically endangered kākāpō reveal major post-glacial and anthropogenic effects on neutral genetic diversity. *Genes*, 9(4): 220.

Dussex N, von Seth J, Knapp M, Kardailsky O, Robertson BC & Dalén L. (2019). Complete genomes of two extinct New Zealand passerines show responses to climate fluctuations but no evidence for genomic erosion prior to extinction. *Biology Letters*, 15: 20190491.

Earl DA & von Holdt BM. (2012). STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, 4(2): 359 – 361.

Edgar GJ, Banks S, Fariña JM, Calvopiña M & Martínez C. (2004). Regional biogeography of shallow reef fish and macro-invertebrate communities in the Galapagos archipelago. *Journal of Biogeography*, 31(7): 1107 – 1124.

Elshire RJ, Glaubitz JC, Sun Q, Poland JA, Kawamoto K, Buckler ES & Mitchell SE. (2011). A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PloS One*, 6(5): e19379.

Emerson BC. (2002). Evolution on oceanic islands: molecular phylogenetic approaches to understanding pattern and process. *Molecular Ecology*, 11(6): 951 – 966.

Emslie SD, McKenzie A & Patterson WP. (2018). The rise and fall of an ancient Adélie penguin 'supercolony' at Cape Adare, Antarctica. *Royal Society Open Science*, 5(4): 172032.

Esper O & Gersonde R. (2014). New tools for the reconstruction of Pleistocene Antarctic sea ice. *Palaeogeography, Palaeoclimatology and Palaeoecology*, 399(1): 260 – 283.

Evanno G, Regnaut S & Goudet J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology.* 14(8), 2611 – 2620.

Excoffier L. (1995). AMOVA 1.55 (analysis of molecular variance). *University of Geneva, Geneva.*

Excoffier L, Estoup A & Cornuet JM. (2005). Bayesian analysis of an admixture model with mutations and arbitrarily linked markers. *Genetics*, 169(3): 1727 – 1738.

Excoffier L, Dupanloup I, Huerta-Sánchez E, Sousa VC & Foll M. (2013). Robust demographic inference from genomic and SNP data. *PLOS Genetics*. 9, e1003905.

Fagundes NJ, Ray N, Beaumont M, Neuenschwander S, Salzano FM, Bonatto SL & Excoffier L. (2007). Statistical evaluation of alternative models of human evolution. *Proceedings of the National Academy of Sciences USA*, 104(45): 17614 – 17619.

Falla RA. (1962). Exploitation of seals, whales and penguins in New Zealand. *New Zealand Ecological Society, Inc*.

Falush D, Stephens M & Pritchard JK. (2003). Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics*, 164(4): 1567 – 1587.

Fleischer RC, McIntosh CE & Tarr CL. (1998). Evolution on a volcanic conveyor belt: using phylogeographic reconstructions and K-Ar-based ages of the Hawaiian Islands to estimate molecular evolutionary rates. *Molecular Ecology*, 7(4): 533 – 545.

Fleming CA. (1979). The geological history of New Zealand and its life. *Auckland University Press*, Auckland.

Forcada J, Trathan PN, Reid K, Murphy EJ & Croxall JP. (2006). Contrasting population changes in sympatric penguin species in association with climate warming. *Global Change Biology*, 12(3): 411 – 423.

Forcada J & Trathan PN. (2009). Penguin responses to climate change in the Southern Ocean. *Global Change Biology*, 15(7): 1618 – 1630.

Fordyce RE & Jones CM. (1990). Penguin history and new fossil material from New Zealand. Penguin Biology. *Academic Press*, San Diego.

Fraser CI, Nikula R, Spencer HG & Waters JM. (2009). Kelp genes reveal effects of sub-Antarctic sea ice during the Last Glacial Maximum. *Proceedings of the National Academy of Sciences USA*, 106(9): 3249 – 3253.

Fraser CI, Nikula R, Ruzzante DE & Waters JM. (2011). Poleward bound: biological impacts of Southern Hemisphere glaciation. *Trends in Ecology and Evolution*, 27(8): 462 – 471.

Fraser CI, Zuccarello GC, Spencer HG, Salvatore LC, Garcia GR & Waters JM. (2013). Genetic affinities between trans-oceanic populations of non-buoyant macroalgae in the high latitudes of the Southern Hemisphere. *PLoS One*, 8(7): e69138.

Fraser CI, Morrison AK, Hogg A, Macaya EC, van Sebille E, Ryan PG, Padovan A, Jack C, Valdivia N & Waters JM. (2018). Antarctica's ecological isolation will be broken by storm-driven dispersal and warming. *Nature Climate Change*, 8(8): 604 – 708.

Fretwell PT & Trathan PN. (2019). Emperors on thin ice: three years of breeding failure at Halley Bay. *Antarctic Science,* 31(3): 133 – 138.

Frenot Y, Gloaguen JC, Massé L & Lebouvier M. (2001). Human activities, ecosystem disturbance and plant invasions in subantarctic Crozet, Kerguelen and Amsterdam Islands. *Biological Conservation*, 101(1): 33 – 50.

Frenot Y, Chown SL, Whinam J, Selkirk PM, Convey P, Skotnicki M & Bergstrom DM. (2004). Biological invasions in the Antarctic: extent, impacts and implications. *Biological Reviews*, 80(1): 45 – 72.

Friesen VL, Burg TM & McCoy KD. (2007). Mechanisms of population differentiation in seabirds. *Molecular Ecology*, 16(9): 1765 – 1785.

Friesen VL. (2015). Speciation in seabirds: why are there so many species… and why aren't there more? *Journal of Ornithology*, 156(1): 27 – 39.

Frost PGH, Siegfried WR & Cooper J. (1976). Conservation of the jackass penguin (*Spheniscus demersus* (L.)). *Biological Conservation*, 9(2): 79 – 99.

Frugone M-J, Lowther A, Noll D, Ramos B, Pistorius P, Dantas GPM, Petry MV, Bonadonna F, Steinfurth A, Polanowski A, Raya Rey A, Lois NA, Pütz K, Trathan P, Wienecke B, Poulin E & Vianna JA. (2018). Contrasting phylogeographic pattern among *Eudyptes* penguins around the Southern Ocean. *Nature Scientific Reports*, 8(1): 17481.

Frugone M-J, López ME, Segovia N, Cole TL, Lowther A, Pistorius P, Dantas G, Petry MV, Bonadonna F, Trathan P, Polanowski A, Wienecke B, Bi K, Wang-Claypool C, Waters JM, Bowie RCK, Poulin E & Vianna JA. (2019). More than the eye can see: Genomic insights into the drivers of genetic differentiation in Royal/Macaroni penguins across the Southern Ocean. *Molecular Phylogenetics and Evolution.* 139: 106563.

Gamble JA & Morris PA. (1989). Sub-Antarctic and Chatham Islands. In Intraplate volcanism in eastern Australia and New Zealand. *Cambridge University Press*, Cambridge.

Gaston KJ, Jones AG, Hänel C & Chown SL (2003). Rates of species introduction to a remote oceanic island. *Proceedings of the Royal Society London B: Biological Sciences*, 270(1519): 1091 – 1098.

Gathorne-Hardy FJ & Jones DT. (2000). The recolonization of the Krakatau islands by termites (Isoptera), and their biogeographical origins. *Biological Journal of the Linnean Society*, 71(2): 251 – 267.

Gavryushkina A, Heath TA, Ksepka DT, Stadler T, Welch D & Drummond AJ. (2017). Bayesian total-evidence dating reveals the recent crown radiation of penguins. *Systematic Biology*, 66(1): 57–73.

Gersonde R, Crosta X, Abelmann A & Armand L. (2005). Sea-surface temperature and sea ice distribution of the Southern Ocean at the EPILOG Last Glacial Maximum – a circum-Antarctic view based on siliceous microfossil records. *Quaternary Science Reviews*, 24(7 – 9): 869 – 896.

Gilbert KJ, Andrew RL, Bock DG, Franklin MT, Kane NC, Moore J-S, Moyers BT, Renaut S, Rennison DJ, Veen T & Vines TH. (2012). Recommendations for utilizing and reporting population genetic analyses: the reproducibility of genetic clustering using the program STRUCTURE. *Molecular Ecology*, 21(20): 4925 – 4930.

Gilbert MTP, Bandelt H-J, Hofreiter M & Barnes I. (2005). Assessing ancient DNA studies. *Trends in Ecology and Evolution*, 20(10): 541 – 554.

Gill BJ. (2010). Regional comparisons of the thickness of moa eggshell fragments (Aves: Dinornithiformes). *Records of the Australian Museum*, 62(1): 115 – 122.

Gillespie R. (2004). Community assembly through adaptive radiation in Hawaiian spiders. *Science*, 303(5656): 356 – 359.

Gillespie RG, Baldwin BG, Waters JM, Fraser CI, Nikula R & Roderick GK. (2012) Long-distance dispersal: a framework for hypothesis testing. *Trends in Ecology and Evolution*, 27(1): 47 – 56.

Godoy E, Marquardt C & Blanco N. (2003). Carta Caldera: Región de Atacama. Servicio Nacional de Geología y Minería, Carta Geológica de Chile, Serie Geología Básica, escala 1:100.000, Santiago.

Göhlich UB. (2007). The oldest fossil record of the extant penguin genus *Spheniscus—a ne*w species from the Miocene of Peru. *Acta Palaeontologica Polonica*, 52(2): 285 – 298.

Gonen S, Bishop SC & Houston RD. (2015). Exploring the utility of cross-laboratory RAD-sequencing datasets for phylogenomic analyses. *BMC Research Notes*, 8(1): 299.

González-Wevar CA, Nakano T, Cañete JI & Poulin E. (2011). Concerted genetic, morphological and ecological diversification in Nacella limpets in the Magellanic Province. *Molecular Ecology*, 20(9): 1936 – 51.

González-Wevar CA, Hüne M, Cañete JI, Mansilla A, Nakano T & Poulin E. (2012). Towards a model of postglacial biogeography in shallow marine species along the Patagonian Province: lessons from the limpet *Nacella magellanica* (Gmelin, 1791). *BMC Evolutionary Biology*, 2(1): 139.

González-Wevar CA, Saucède T, Morley SA, Chown SL & Poulin E. (2013). Extinction and recolonization of maritime Antarctica in the limpet *Nacella concinna* (Strebel, 1908) during the last glacial cycle: toward a model of Quaternary biogeography in shallow Antarctic invertebrates. *Molecular Ecology*, 22(20): 5221 – 36.

González-Wevar CA, Chown SL, Morley S, Coria N, Saucéde T & Poulin E. (2016). Out of Antarctica: quaternary colonization of sub-Antarctic Marion Island by the limpet genus *Nacella* (Patellogastropoda: Nacellidae). *Polar Biology*, 39(1): 77 – 89.

Gradstein FM. (2012). The Geologic Time Scale 2012. *Elsevier*, Amsterdam.

Grant PR & Grant BR. (2002). Adaptive radiation of Darwin's finches: Recent data help explain how this famous group of Galápagos birds evolved, although gaps in our understanding remain. *American Scientist*, 90(2): 130 – 139.

Grealy AC, McDowell MC, Scofield RP, Murray DC, Fusco DA, Haile J, Prideaux GJ & Bunce M. (2015). A critical evaluation of how ancient DNA bulk bone metabarcoding complements traditional morphological analysis of fossil assemblages. *Quaternary Science Reviews*, 128: 37 – 47.

Grealy A, Bunce M & Holleley C. (2019). Avian mitochondrial genomes retrieved from museum eggshell. *Molecular Ecology Resources*. 19(4): 1052 – 1062.

Green RE, Krause J, Ptak SE, Briggs AW, Ronan MT, Simons JF, Du L, Egholm M, Rothberg JM, Paunovic M & Pääbo S. (2006). Analysis of one million base pairs of Neanderthal DNA. *Nature*, 444(7117): 330 – 336.

Groot M, Ervynck A & Pigière F. (2010). Vagrant vultures: archaeological evidence for the cinereous vulture (*Aegypius monachus*) in the Low Countries, in Birds in Archaeology. *Proceedings of the 6th Meeting of the ICAZ Bird Working Group in Groningen (23.8 – 27.8.2008)*, Groningen.

Grosser S, Burridge CP, Peucker AJ & Waters JM. (2015). Coalescent modelling suggests recent secondary-contact of cryptic penguin species. *PLoS One*, 10(12): e0144966.

Grosser S, Rawlence NJ, Anderson CNK, Smith IWG, Scofield RP & Waters JM. (2016). Invader or resident? Ancient-DNA reveals rapid species turnover in New Zealand little penguins. *Proceedings of the Royal Society B: Biological Sciences*, 283(1824): 20152879.

Grosser S, Scofield RP & Waters JM. (2017). Multivariate skeletal analyses support a taxonomic distinction between New Zealand and Australian *Eudyptula* penguins (Sphenisciformes: Spheniscidae). *Emu – Austral Ornithology* 177(3): 176 – 283.

Gruber B, Unmack P, Berry O & Georges A. (2017). DARTR: An R package to facilitate analysis of SNP data generated from reduced representation genome sequencing. *Molecular Ecology Resources*, 18(3), 691 – 699.

Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W & Gascuel O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Systematic Biology*, 59(3): 307 – 321.

Gutenkunst RN, Hernandez RD, Williamson SH & Bustamante CD. (2009). Inferring the Joint Demographic History of Multiple Populations from Multidimensional SNP Frequency Data. *PLOS Genetics*. 5: e1000695.

Hays C. (1985). The Humboldt penguin in Peru. *Oryx*, 18(2): 92 – 95.

Hasegawa M, Kishino H & Yano T. (1985). Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *Journal of Molecular Evolution*, 22(2): 160 – 174.

Hebert PDN, Stoeckle MY, Zemlak TS & Francis CM. (2004). Identification of Birds through DNA Barcodes. *PLoS Biology*, 2(10): e312.

Heenan PB, Mitchell AD, De Lange PJ, Keeling J & Paterson AM. (2010). Late-Cenozoic origin and diversification of Chatham Islands endemic plant species revealed by analyses of DNA sequence data. *New Zealand Journal of Botany*, 48(2): 83 – 136.

Heerah K, Dias MP, Delord K, Oppel S, Barbraud C, Weimerskirch H & Bost CA. (2019). Important areas and conservation sites for a community of globally threatened marine predators of the Southern Indian Ocean. *Biological Conservation*, 234(1): 192 – 201.

Heintzman PD, Soares AER, Chang D & Shapiro B. (2015). Paleogenomics. *Reviews in Cell Biology and Molecular Medicine*. 1: 243 – 267.

Hewitt G. (2000). The genetic legacy of the Quaternary ice ages. *Nature*, 405(6789): 907 – 913.

Hewitt GM. (2004). Genetic consequences of climatic oscillations in the Quaternary. *Philosophical Transactions of the Royal Society of London B Biological Series.* 359(1442): 183 – 195.

Hindwood KA. (1938). The occurrence of crested penguins in Australian waters: with particular reference to *Eudyptes pachyrhynchus*. *Emu*, 38(4): 377 – 379.

Hiscock JA & Chilvers LB. (2014). Declining eastern rockhopper (*Eudyptes filholi*) and erect-crested (*E. sclateri*) penguins on the Antipodes Islands, New Zealand. *New Zealand Journal of Ecology*, 38(1): 124 – 131.

Hiscock JA & Chilvers LB. (2016). Snares crested penguins *Eudyptes robustus* population estimates 2000 – 2013. *New Zealand Journal of Ecology*, 40(1): 108 – 113.

Ho SYW, Kolokotronis SO & Allaby RG. (2007). Elevated substitution rates estimated from ancient DNA sequences. *Biology Letters*, 3(6): 702 – 705.

Ho SYW, Lanfear R, Phillips MJ, Barnes I, Thomas JA, Kolokotronis S-O & Shapiro B. (2011). Bayesian estimation of substitution rates from ancient DNA sequences with low information content. *Systematic Biology*, 60(3): 1 – 10.

Hodgson DA, Graham AGC, Roberts SJ, Bentley MJ, Cofaigh Có, Verleyen E, Vyverman W, Jomelli V, Favier V, Brunstein D, Verfaillie D, Colhoun EA, Saunders KM, Selkirk PM, Mackintosh A, Dedding DW, Nel W, Hall K, McGlone MS, Van der Putten N, Dickens WA & Smith JA. (2014). Terrestrial and submarine evidence for the extent and timing of the Last Glacial Maximum and the onset of deglaciation on the maritime Antarctic and sub-Antarctic islands. *Quaternary Science Reviews*, 100: 137 – 158.

Hoffman EA & Blouin MS. (2004). Historical data refute recent range contraction as cause of low genetic diversity in isolated frog populations. *Molecular Ecology*, 13(2): 271 – 276.

Hofman RJ. (2017). Sealing, whaling and krill fishing in the Southern Ocean: past and possible future effects on catch regulations. *Polar Record*, 53(1): 88 – 99.

Hogg AG, Blackwell PG, Niu M, Buck CE, Guilderson TP, Heaton TJ, Palmer JG, Reimer PJ, Reimer RW, Turney CSM & Zimmerman SRH. (2013). SHCal13 Southern Hemisphere Calibration, 0 – 50,000 Years cal BP. *Radiocarbon* 55(4): 1889 – 1903.

Holm S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics,* pp. 65 – 70.

Holm-Hansen O. (1985). Nutrient cycles in Antarctic marine ecosystems. In Antarctic nutrient cycles and food webs. *Springer*, Heidelberg.

Hume JP. (2017). Extinct Birds: Second Edition. *Bloomsbury Publishing Plc*, London.

Iredale T & Cayley NW. (1925). Australian crested penguins. *Emu-Austral Ornithology*, 25(1): 1 – 6.

Jadwiszczak P. (2006). Eocene penguins of Seymour Island, Antarctica: the earliest record, taxonomic problems and some evolutionary considerations. *Polish Polar Research*, 27(4): 287 – 302.

Jagatap H, Monsanto DM, van Vuuren BJ, Parbhu SP, Dinoi A, Janion-Scheepers C, Sekar S, Teske PR & Emami-Khoyi A. (2019). The complete mitogenome of *Isotomurus maculatus*: a widespread species that is invading the sub-Antarctic region. *Mitochondrial DNA Part B*, 4(1): 1706 – 1708.

Janes JK, Miller JM, Dupuis JR, Malenfant RM, Gorrell JC, Cullingham CI & Andrew RL. (2017). The K = 2 conundrum. *Molecular Ecology*, 26(14): 3594 – 3602.

Jarvis ED, Mirarab S, Aberer AJ, Li B, Houde P, Li C, Ho SYW, Faircloth BC, Nabholz B, Howard JT, Suh A, Weber CC, da Fonseca RR, Li J, Zhang F, Li H, Zhou L, Narula N, Liu L, Ganapathy G, Boussau B, Bayzid MS, Zavidovych V, Subramanian S, Gabaldón T, Capella-Gutiérrez S, Huerta-Cepas J, Rekepalli B, Munch K, Schierup M, Lindow B, Warren WC, Ray D, Green RE, Bruford MW, Zhan X, Dixon A, Li S, Li N, Huang Y, Derryberry EP, Bertelsen MF, Sheldon FH, Brumfield RT, Mello CV,

Lovell PV, Wirthlin M, Cruz Schneider MP, Prosdocimi F, Samaniego JA, Vargas Velazquez AM, Alfaro-Núñez A, Campos PF, Petersen B, Sicheritz-Ponten T, Pas A, Bailey T, Scofield P, Bunce M, Lambert DM, Zhou Q, Perelman P, Driskell AC, Shapiro B, Ziong Z, Zeng Y, Liu S, Li Z, Liu B, Wu K, Xiao J, Yingi X, Zheng Q, Zhang Y, Yang H, Wang J, Smeds L, Rheindt FE, Braun M, Fjeldsa J, Orlando L, Barker FK, Jønsson KA, Johnson W, Koepfli K-P, O'Brien S, Haussler D, Ryder OA, Rahbek C, Willerslev E, Graves GR, Glenn TC, McCormack J, Burt D, Ellegren H, Alström P, Edwards SV, Stamatakis A, Mindell DP, Cracraft J, Braun EL, Warnow T, Jun W, Gilbert MTP & Zhang G. (2014). Whole genome analyses resolve early branches in the tree of life of modern birds. *Science*, 346(6215): 1320 – 1331.

Jarvis ED Mirarab S, Aberer AJ, Li B, Houde P, Li C, Ho SYW, Faircloth BC, Nabholz B, Howard JT, Suh A, Weber CC, da Fonseca RR, Alfari-Núñez-Alfaro A, Narula N, Liu L, Burt D, Ellegren H, Edwards SV, Stamatakis A, Mindell DP, Cracraft J, Braun EL, Warnow T, Jun W, Gilbert MTP, Zhang G & The Avian Phylogenomics Consortium. (2015). Phylogenomic analyses data of the avian phylogenomics project. *GigaScience*, 4(1): 4.

Jenouvrier S, Caswell H, Barbraud C, Holland M, Strœve J & Weimerskirch H. (2009). Demographic models and IPCC climate projections predict the decline of an emperor penguin population. *Proceedings of the National Academy of Sciences USA*, 106(6): 1844 – 1847.

Jenouvrier S, Holland M, Stroeve J, Serreze M, Barbraud C, Weimerskirch H & Caswell H. (2014). Projected continent-wide declines of the emperor penguin under climate change. *Nature Climate Change*, 4(8): 715 – 718.

Jombart T & Collins C. (2015). Analysing genome-wide SNP data using adegenet 2.0.0. <http://adegenet.r-forge.r-project.org/files/tutorial-genomics.pdf>. Last accessed 04 September 2019.

Jouventin P, Cuthbert RJ & Ottvall R. (2006). Genetic isolation and divergence in sexual traits: evidence for the northern rockhopper penguin *Eudyptes moseleyi* being a sibling species. *Molecular Ecology*, 15(11): 3413 – 3423.

Jouventin P & Dobson FS. (2017). Why penguins communicate: The evolution of visual and vocal signals. Academic Press.

Jouzel J, Masson-Delmotte V, Cattani O, Dreyfus G, Falourd S, Hoffman G, Minster B, Nouet J, Barnola JM, Chappellaz J, Fischer H, Gallet JC, Johnsen S, Leuenberger M, Loulergue L, Luethi D, Oerter H, Parrenin F, Raisbeck G, Raynaud D, Schilt A, Schwander J, Selmo E, Souchez R, Spahni R, Stauffer B, Steffensen JP, Stenni B, Stocker TF, Tison JL, Werner M & Wolff EW. (2007). Orbital and millennial Antarctic climate variability over the past 800,000 years. *Science*, 317(5839): 793 – 796.

Jónsson H, Ginolhac A, Schubert M, Johnson P & Orlando L. (2013). mapDamage 2.0: fast approximate Bayesian estimates of ancient DNA damage parameters. *Bioinformatics*, 29(13): 1682 – 1684.

Juan C, Emerson BC, Oromí P & Hewitt GM. (2000). Colonization and diversification: towards a phylogeographic synthesis for the Canary Islands. *Trends in Ecology and Evolution*, 15(3): 104 – 109.

Kamvar ZN, Tabima JF & Grünwald NJ. (2014). Poppr: an R package for genetic analysis of populations with clonal, partially clonal, and or/sexual reproduction. *PeerJ*, 2: e281.

Kamvar ZN, Brooks JC & Grünwald NJ. (2015). Novel R tools for analysis of genome-wide population genetic data with emphasis on clonality. *Frontiers of Genetics*, 6: 208.

Kearse M, Wilson MR, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, Thierer T, Ashton B, Meintjes P & Drummond A. (2012). Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*, 28(2): 1647 – 1649.

Kemper J, Roux JP, Bartlett PA, Chesselet YJ, James JAC, Jones R, Wepener S & Molloy FJ. (2001). Recent population trends of African penguins *Spheniscus demersus* in Namibia. *African Journal of Marine Science,* 23(1): 429 – 434.

Keogan K, Daunt F, Wanless S, Phillips RA, Walling CA, Agnew P, Ainley DG, Anker-Nilssen T, Ballard G, Barrett RT, Barton KJ, Bech C, Becker P, Berglund P-A, Bollache L, Bond AL, Bouwhuis S, Bradley RW, Burr ZM, Camphuysen K, Catry P, Chiaradia A, Christensen-Dalsgaard S, Cuthbert R, Dehnhard N, Descamps S, Diamond T, Divoky G, Drummond H, Dugger KM, Dunn MJ, Emmerson L, Erikstad KE, Fort J, Fraser W, Genovart M, Gilg O, González-Solís J, Granadeiro JP, Grémillet D, Hansen J, Hanssen SA, Harris M, Hedd A, Hinke J, Igual JM, Jahncke J, Jones I, Kappes PJ, Lang J, Langset M, Lescroël A, Lorentsen S-H, Lyver POB, Mallory M, Moe B, Montevecchi WA, Monticelli D, Mostello C, Newell M, Nicholson L, Nisbet I, Olsson O, Oro D, Pattison V, Poisbleau M, Pyk T, Quintana F, Ramos JA, Ramos R, Reiertsen TK, Rodríguez C, Ryan P, Sanz-Aguilar A, Schmidt NM, Shannon P, Sittler B, Southwell C, Surman C, Svageli WS, Trivelpiece W, Warzybok P, Watanuki Y, Weimerskirch H, Wilson PR, Wood AG, Phillimore AB & Lewis S. (2018). Global phenological insensitivity to shifting ocean temperatures among seabirds. *Nature Climate Change*, 8(4): 313 – 318.

Kier G, Kreft H, Lee TM, Jetz W, Ibisch PL, Nowicki C, Mutke J & Barthlott W. (2009). A global assessment of endemism and species richness across island and mainland regions. *Proceedings of the National Academy of Sciences USA*, 106(23): 9322 – 9327.

Kilian A, Wenzl P, Huttner E, Carling J, Xia L, Blois H, Caig V, Heller-Uszynska K, Jaccoud D, Hopper C, Aschenbrenner-Kilian M, Evers M, Peng K, Cayla C, Hok P & Uszynski G. (2012). Diversity Arrays Technology: a generic genome profiling technology on open platforms. *Methods in Molecular Biology*, 888: 67 – 89.

Kimball RT, Oliveros CH, Wang N, White ND, Barker FK, Field DJ, Ksepka DT, Chesser RT, Moyle RG, Braun MJ, Brumfield RT, Faircloth BC, Tilston Smith B & Braun EL. (2019). A phylogenomic supertree of birds. *Diversity*, 11(7): 109.

King SD, Harper GA, Wright JB, McInnes JC, van der Lubbe JE, Dobbins ML & Murray SJ. (2012). Site-specific reproductive failure and decline of a population of the Endangered yellow-eyed penguin: a case for foraging habitat quality. *Marine Ecology Progress Series*, 467(1): 233 – 244.

Kinsky FC. (1970). Annotated checklist of the birds of New Zealand including the birds of the Ross Dependency. *Reed*, Wellington.

Kircher M. (2012). Analysis of high-throughput ancient DNA sequencing Data. In Ancient DNA: Methods and Protocols. University of Hertfordshire, Hertfordshire.

Knapp M, Clarke AC, Horsburgh KA & Matisoo-Smith EA. (2012). Setting the stage – Building and working in an ancient DNA laboratory. *Annals of Anatomy – Anatomischer Anzeiger*, 194(1): 3 – 6.

Knapp M, Thomas JE, Haile J, Prost S, Ho SYW, Dussex N, Cameron-Christie S, Kardailsky O, Barnett R, Bunce M, Gilbert MTP & Scofield RP. (2019). Mitogenomic evidence of close relationships between New Zealand's extinct giant raptors and small-sized Australian sister-taxa. *Molecular Phylogenetics and Evolution*. 134: 122 – 128.

Krogh A, Brown M, Mian IS, Sjölander K & Haussler D. (1994). Hidden Markov models in computational biology: Applications to protein modeling. *Journal of Molecular Biology*, 235(5): 1501 – 1531.

Ksepka DT, Bertelli S & Giannini NP. (2006). The phylogeny of the living and fossil Sphenisciformes (penguins). *Cladistics*, 22(5): 412 – 441.

Ksepka D & Clarke JA. (2010). The basal penguin (Aves: Sphenisciformes) *Perudyptes devriesi* and a phylogenetic evaluation of the penguin fossil record. *Bulletin of the American Museum of Natural History*. 337: 77.

Ksepka DT & Ando T. (2011). Penguins past, present, and future: trends in the evolution of the Sphenisciformes. *Living Dinosaurs*, pp. 155 – 186.

Ksepka DT & Thomas DB. (2012). Multiple cenozoic invasions of Africa by penguins (Aves, Sphenisciformes). *Proceedings of the Royal Society B: Biological Sciences*, 279(1730): 1027 – 1032.

Ksepka DT, Fordyce RE, Ando T & Jones CM. (2012). New fossil penguins (Aves: Spheniciformes) from the Oligocene of New Zealand reveal the skeletal plan of stem penguins. *Journal of Vertebrate Paleontology*, 32(2): 235 – 254.

Ksepka DT & Phillips MJ. (2015). Avian diversification patterns across the K-Pg boundary: Influence of calibrations, datasets, and model misspecification. *Annals of the Missouri Botanical Garden*, 100(4): 300 – 328.

Lalas C, Hamel J, Tennyson AJD & Worthy TH. (2014). Southern extensions for Holocene records of Australian pelican (*Pelecanus conspicillatus*) and New Zealand musk duck (*Biziura delautouri*) in New Zealand. *Notornis*, 61(2): 106 – 108.

Lanfear R, Calcott B, Ho SY & Guindon S. (2012). PartitionFinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. *Molecular Biology and Evolution*, 29(6): 1695 – 1701.

Lanfear R, Frandsen PB, Wright AM, Senfeld T & Calcott B. (2016). PartitionFinder 2: new methods for selecting partitioned models of evolution from molecular and morphological phylogenetic analyses. *Molecular Biology and Evolution*, 34(3): 772 – 773.

Leach BF. (1979). In Birds of a feather: Osteological and archaeological papers from the South Pacific in honour of R. J. Scarlett. *New Zealand Archaeological Association Monograph II*. Oxford.

Le Bohec C, Durant JM, Gauthier-Clerc M, Stenseth NC, Park Y-H, Pradel R, Grémillet D, Gendner J-P & Le Maho Y. (2008). King penguin population threatened by Southern Ocean warming. *Proceedings of the National Academy of Sciences USA,* 105(7): 2493 – 2497.

Lees AC & Gilroy JJ. (2009). Vagrancy Mechanisms in Passerines and Near-Passerines. In: Slack R. Rare Birds, Where and When: An analysis of status and distribution in Britain and Ireland. Volume 1: sandgrouse to New World orioles. *Rare Bird Books*, York.

Leigh JW & Bryant D. (2015). PopART: full-feature software for haplotype network construction. *Methods in Ecology and Evolution*, 6(9): 1110 – 1116.

Lerner HR, Meyer M, James HF, Hofreiter M & Fleischer RC. (2011). Multilocus resolution of phylogeny and timescale in the extant adaptive radiation of Hawaiian honeycreepers. *Current Biology*, 21(21): 1838 – 1844.

Lescroël A & Bost C-A. (2005). Foraging under contrasting oceanographic conditions: the gentoo penguin at Kerguelen Archipelago. *Marine Ecology Progress Series*, 302: 245 – 261.

Levy H, Clucas G, Rogers AD, Polito MJ, Lynch H, Emslie SD & Hart T. (2016). Population structure and phylogeography of the Gentoo Penguin (*Pygoscelis papua*) across the Scotia Arc. *Ecology and Evolution*, 6(6): 1834 – 1853.

Lewin R. (1985). Hawaiian Drosophila: Young islands, old flies. *Science*, 229(4718): 1072 – 1074.

Li H & Durbin R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, 25(14): 1754 – 1760.

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R & 1000 Genome Project Data Processing Subgroup. (2009) The Sequence Alignment/Map (SAM) format and SAMtools. *Bioinformatics*, 25(16): 2078 – 2079.

Li C, Zhang Y, Li J, Kong L, Hu H, Pan H, Xu L, Deng Y, Li Q, Jin L, Yu H, Chen Y, Liu B, Yang L, Liu S, Zhang Y, Lang Y, Xia J, He W, Shi Q, Subramanian S, Millar CD, Meader S, Rands CM, Fujita MF, Greenwold MJ, Castoe TA, Pollock DD, Gu W, Nam K, Ellegren H, Ho SYW, Burt DW, Ponting CP, Jarvis ED, Gilbert MTP, Yang H, Wang J, Lambert DM, Wang J & Zhang G. (2014). Two Antarctic penguin genomes reveal insights into their evolutionary history and molecular changes related to the Antarctic environment. *GigaScience*, 3(1): 27.

Lischer HEL & Excoffier L. (2012). PGDSpider: An automated data conversion tool for connecting population genetics and genomics programs. *Bioinformatics*, 28(2): 298 – 299.

Lisiecki LE & Raymo ME. (2005). A Pliocene-Pleistocene stack of globally distributed benthic δ18O records. *Paleoceanography*, 20: PA1003.

Lomolino MV. (2005). Body size evolution in insular vertebrates: generality of the island rule. *Journal of Biogeography*, 32(10): 1683 – 1699.

Loope LL, Hamann O & Stone CP. (1988). Comparative conservation biology of oceanic archipelagos: Hawaii and the Galápagos. *BioScience*, 38(4): 272 – 282.

Lopes JS & Boessenkool S. (2010). The use of approximate Bayesian computation in conservation genetics and its application in a case study on yellow-eyed penguins. *Conservation Genetics*, 11(2): 421 – 433.

Losos JB, Jackman TR, Larson A, de Queiroz K & Rodríguez-Schettino L. (1998). Contingency and determinism in replicated adaptive radiations of island lizards. *Science*, 279(5359): 2115 – 2118.

Losos JB & Ricklefs RE. (2009). Adaptation and diversification on islands. *Nature*, 457(7231): 830.

Lynch HJ, Naveen R, Trathan PN & Fagan WF. (2012). Spatially integrated assessment reveals widespread changes in penguin populations on the Antarctic Peninsula. *Ecology*, 93(6): 1367 – 1377.

MacArthur RH. & Wilson EO. (1967). The Theory of Island Biogeography. *Princeton University Press*, Princeton.

Maggs CA, Castilho R, Foltz D, Henzler C, Jolly MT, Kelly J, Olsen J, Perez KE, Stam W, Väinölä R, Viard F & Wares J. (2008). Evaluating signatures of glacial refugia for North Atlantic benthic marine taxa. *Ecology*, 89(Sp11): S108 – S122.

Marchant S & Higgins PJ. (1990). Handbook of Australian, New Zealand and Antarctic birds. Vol. 1, Pt. B. *Oxford University Press*, Melbourne.

Marquardt C, Blanco N, Godoy E, Lavenu A, Ortlieb L, Marchant M & Guzmán N. (2000). Estratigrafía del Cenozoico Superior en el área de Caldera (26°45' – 28°S), III Región de Atacama. In Congreso Geológico Chileno, No. 9, Actas: 504 – 508. Puerto Varas.

Marshall J. & Speer K. (2012). Closure of the meridional overturning circulation through Southern Ocean upwelling. *Nature Geoscience*, 5(3): 171.

Marske KA, Rahbek C & Nogués-Bravo D. (2013). Phylogeography: spanning the ecology-evolution continuum. *Ecography*, 36(11): 1169 – 1181.

Martínez I. (1992). In Handbook of the birds of the world, Ostrich to Ducks Vol. 1. Lynxs editions.

Matocq MD & Villablanca FX. (2001). Low genetic diversity in an endangered species: recent or historic pattern? *Biological Conservation*, 98(1): 61 – 68.

Mattern T. (2013). Snares penguin (*Eudyptes robustus*). In Penguins: natural history and conservation. Washington University Press, Washington.

Mattern T & Long R. (2017). Survey and population size estimate of Fiordland penguin (tawaki; *Eudyptes pachyrhynchus*) in Milford Sound/Piopiotahi, New Zealand. *Notornis* 64(2): 97 – 101.

Mattern T, Meyer S, Ellenberg U, Houston DM, Darby JT, Young M, van Heezik Y & Seddon PJ. (2017). Quantifying climate change impacts emphasises the importance of managing regional threats in the endangered Yellow-eyed penguin. *PeerJ*, 5: e3272.

Mattern T, Pütz K, Garcia-Borboroglu P, Ellenberg U, Houston DM, Long R, Lüthi B & Seddon PJ. (2018). Marathon penguins – Reasons and consequences of long-range dispersal in Fiordland penguins/Tawaki during the pre-moult period. *PloS One*, 13(8): e0198688.

Mattern T & Wilson K-J. (2018). New Zealand penguins – current knowledge and research priorities. A report compiled for Birds New Zealand, July 2018.

Maund JG, Rex DC, Le Roex AP & Reid DL. (1988). Volcanism on Gough Island: a revised stratigraphy. *Geological Magazine*, 125(2): 175 – 181.

Maxwell JJ & Smith IWG. (2015). A reassessment of settlement patterns and subsistence at Point Durham, Chatham Island. *Archaeology in Oceania*, 50(3): 162 – 174.

Mayr G, Scofield RP, De Pietri VL & Tennyson AJD. (2017). A Paleocene penguin from New Zealand substantiates multiple origins of gigantism in fossil Sphenisciformes. *Nature Communications*, 8(1): 1927.

Mays HL, Oehler DA, Morrison KW, Morales AE, Lycans A, Perdue J, Battley PF, Cherel Y, Chilvers BL, Crofts S, Demongin L, Fry WR, Hiscock J, Kusch A, Marin M, Poisbleau M, Quillfeldt P, Raya Rey A, Steinfurth A, Thompson DR & Weakley LA. (in press). Phylogeography, population structure, and species delimination in rockhopper penguins (*Eudyptes chrysocome* and *Eudyptes moseleyi*). *Journal of Heredity*.

McCulloch G & Waters JM. (in press). Phylogenetic divergence of island biotas: calibration, extinction, and 'relict' lineages. *Molecular Ecology*.

McDougall I & Ollier CD. (1982). Potassium-argon ages from Tristan da Cunha, south Atlantic. *Geological Magazine*, 119(1): 87 – 93.

McNab BK. (1994). Energy conservation and the evolution of flightlessness in birds. *The American Naturalist*, 144(4): 628 – 642.

McWethy DB, Whitlock C, Wilmshurst JM, McGlone MS, Fromont M, Li X, Dieffenbacher-Krall A, Hobbs WO, Fritz SC & Cook ER. (2010). Rapid landscape transformation in South Island, New Zealand, following initial Polynesian settlement. *Proceedings of the National Academy of Sciences USA*, 107(50): 21343 – 21348.

Meirmans PG & Van Tienderen PH. (2004). GENOTYPE and GENODIVE: two programs for the analysis of genetic diversity of asexual organisms. *Molecular Ecology Notes*, 4(4): 792 – 794.

Mendelson TC & Shaw KL. (2005). Rapid speciation in an arthropod. *Nature*, 433: 375 – 376.

Meredith MAM & Sin FY. (1988a). Genetic variation of four populations of the little blue penguin, *Eudyptula minor*. *Heredity*, 60(1): 69.

Meredith MAM & Sin FY. (1988b). Morphometrical analysis of four populations of the Little Blue Penguin, *Eudyptula minor*. *Journal of Natural History*, 22(3): 801 – 809.

Meyer M & Kircher M. (2010). Illumina sequencing library preparation for highly multiplexed target capture and sequencing. Cold Spring Harbor Protocols.

Meyer M, Arsuaga J-L, de Filippo C, Nagel S, Aximu-Petri A, Nickel B, Martínez I, Gracia A, Bermúdez de Castro JM, Carbonell E, Viola B, Kelso J, Prüfer K & Pääbo S. (2016). Nuclear DNA sequences from the Middle Pleistocene Sima de los Huesos hominins. *Nature*, 531(7595): 504.

Millar CD, Dodd A, Anderson J, Gibb GC, Ritchie PA, Baroni C, Woodhams MD, Hendy MD & Lambert DM. (2008). Mutation and evolutionary rates in Adélie penguins from the Antarctic. *PLOS Genetics*, 4(10): e1000209.

Millener PR. (1990). Evolution, extinction and the subfossil record of New Zealand's avifauna. In A flying start. *Random Century*, Auckland.

Millener PR. (1999). The history of the Chatham Islands' bird fauna of the last 7000 years-a chronicle of change and extinction. Proceedings of the 4th International meeting of the Society of Avian Paleontology and Evolution (Washington, D.C., June 1996). *Smithsonian Contributions to Paleobiology*, 89: 85 – 109.

Miller MA, Pfeiffer W & Schwartz T. (2010). Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In Proceedings of the Gateway Computing Environments Workshop (GCE), 14 Nov. 2010, New Orleans, LA pp 1 – 8.

Miller W, Drautz DI, Ratan A, Pusey B, Qi J, Lesk AM, Tomsho LP, Packard MD, Zhao F, Sher A, Tikhonov A, Raney B, Patterson N, Lindblad-Toh K, Lander ES, Knight JR, Irzyk GP, Fredrikson KM, Harkins TT, Sheridan S, Pringle T & Schuster SC. (2008). Sequencing the nuclear genome of the extinct wooly mammoth. *Nature*, 456(7220): 387.

Miskelly CM. (1997). Biological notes on the Western Chain, Snares Islands, 1984 – 85 and 1985 – 86. Conservation Advisory Science Notes No. 144. Department of Conservation, Wellington.

Miskelly CM, Bester AJ & Bell M. (2006). Additions to the Chatham Islands' bird list, with further records of vagrant and colonising bird species. *Notornis*, 53(2): 213 – 228.

Mitchell KJ, Wood JR, Scofield RP, Llamas B & Cooper A. (2014a). Ancient mitochondrial genome reveals unsuspected taxonomic affinity of the extinct Chatham duck (*Pachyanas chathamica*) and resolves divergence times for New Zealand and sub-Antarctic brown teals. *Molecular Phylogenetics and Evolution*, 70: 420 – 428.

Mitchell KJ, Llamas B, Soubrier J, Rawlence NJ, Worthy TH, Wood JR, Lee MSY & Cooper A. (2014b). Ancient DNA reveals elephant birds and kiwi are sister taxa and clarifies ratite bird evolution. *Science*, 344(6186): 898 – 900.

Moodley Y, Masello JF, Cole TL, Calderon L, Munimanda GK, Thali MR, Alderman R, Cuthbert RJ, Marin M, Massaro M, Navarro J, Phillips RA, Ryan PG, Suazo CG, Cherel Y, Weimerskirch H & Quillfelft P. (2015). Evolutionary factors affecting the cross-species utility of newly developed microsatellite markers in seabirds. *Molecular Ecology Resources*, 15(5): 1046 – 1058.

Moon KL, Chown SL. & Fraser CI. (2017). Reconsidering connectivity in the sub-Antarctic. *Biological Reviews*, 92(4): 2164 – 2181.

Moore WS & DeFilippis VR. (1997). The window of taxonomic resolution for phylogenies based on mitochondrial cytochrome b. In: Avian molecular evolution and systematics. *Academic Press*, San Diego.

Morales HE, Pavlova A, Amos N, Major R, Kilian A, Greening C & Sunnucks P. (2018). Concordant divergence of mitogenomes and a mitonuclear gene cluster in bird lineages inhabiting different climates. *Nature Ecology and Evolution*, 2(8): 1258.

Morin PA, Luikart G & Wayne RK. (2004). SNPs in ecology, evolution and conservation. *Trends in Ecology and Evolution*, 19(4): 208 – 216.

Morrison KW & Sagar PM. (2014). First record of interbreeding between a Snares crested (*Eudyptes robustus*) and erect-crested penguin (*E. sclateri*). *Notornis* 61(2): 109 – 112.

Morrison KW, Battley PF, Sagar PM & Thompson DR. (2015). Population dynamics of eastern rockhopper penguins on Campbell Island in relation to sea surface temperature 1942–2012: current warming hiatus pauses a long-term decline. *Polar Biology*, 38(2): 163 – 177.

Mortimer N, Kohn B, Seward D, Spell T & Tulloch A. (2015). Reconnaissance thermochronology of southern Zealandia. *Journal of the Geological Society*, 173(2): 370 – 383.

Munro KJ & Burg TM. (2017). A review of historical and contemporary processes affecting population genetic structure of Southern Ocean seabirds. *Emu – Austral Ornithology*, 117(1): 4 – 18.

Mura-Jornet I, Pimentel C, Dantas GPM, Petry MV, González-Acuña D, Barbosa A, Lowther AD, Kovacs KM, Poulin E & Vianna JA. (2018). Chinstrap penguin population genetic structure: one or more populations along the Southern Ocean? *BMC Evolutionary Biology*, 18(1): 90.

Murray DC, Haile J, Dortch J, White NE, Haouchar D, Bellgard MI, Allcock RJ, Prideaux GJ & Bunce M. (2013). Scrapheap Challenge: A novel bulk-bone metabarcoding method to investigate ancient DNA in faunal assemblages. *Scientific Reports*, 3(1): 3371.

Myers N, Mittermeier RA, Mittermeier CG, da Fonseca GAB & Kent J. (2000). Biodiversity hotspots for conservation priorities. *Nature*, 403(6772): 853.

Myrcha A, Tatur A & Del Valle R. (1990). A new species of fossil penguin from Seymour Island, West Antarctica. *Alcheringa*, 14(3): 195 – 200.

Myrcha A. (2006). Eocene penguins of Seymour Island, Antarctica: Taxonomy. *Polish Polar Research*. 27(1): 3 – 62.

Naish TR, Wehland F, Wilson GS, Browne GH, Cook RA, Morgans HEG, Rosenberg M, King PR, Smale D, Nelson CS, Kamp PJJ & Richetts B. (2005). An integrated sequence stratigraphic, palaeoenvironmental, and chronostratigraphic analysis of the Tangahoe Formation, southern Taranaki coast, with implications for mid-Pliocene (c. 3.4 – 3.0 Ma) glacio-eustatic sea-level changes. *Journal of the Royal Society of New Zealand*, 35(1 – 2): 151 – 196.

Napier R. (1968). Erect-crested and rockhopper penguins interbreeding in the Flkland Islands. *British Antarctic Survey*, 16: 71 – 72.

Nei M. (1987). Molecular evolutionary genetics. Columbia University Press.

Nikula R, Fraser CI, Spencer HG & Waters JM. (2010). Circumpolar dispersal by rafting in two sub-Antarctic kelp-dwelling crustaceans. *Marine Ecology Progress Series*, 405: 221 – 230.

Nims BD, Vargas FH, Merkel J & Parker PG. (2008). Low genetic diversity and lack of population structure in the endangered Galápagos penguin (*Spheniscus mendiculus*). *Conservation Genetics*, 9(6): 1413 – 1420.

Nogué S, de Nascimento L, Froyd CA, Wilmshurst JM, de Boer EJ, Coffey EED, Whittaker RJ, Fernández-Palacios JM, Willis KJ. (2017). Island biodiversity conservation needs palaeoecology. *Nature Ecology and Evolution*, 1(7): 0181.

Noonan JP, Coop G, Kudaravalli S, Smith D, Krause J, Alessi J, Chen F, Platt D, Pääbo S, Pritchard JK & Rubin EM. (2006). Sequencing and analysis of Neanderthal genomic DNA. *Science*, 314(5802): 1113 – 1118.

Ogg J, Ogg GG & Gradstein FM. (2008). The Concise Geologic Time Scale. *Cambridge University Press*, Cambridge.

Oliveros CH, Field DJ, Ksepka DT, Barker K, Aleixo A, Andersen MJ, Alström P, Benz BW, Braun EL, Braun MJ, Bravo GA, Brumfield RT, Chesser RT, Claramunt S, Cracraft J, Cuervo AM, Derryberry EP, Glenn TC, Harvey MG, Hosner PA, Joseph L, Kimball RT, Mack AL, Miskelly CM, Peterson AT, Robbins MB, Sheldon FH, Silveirau LF, Smith BT, White ND, Moyle RG & Faircloth BC. (2019). Earth history and the passerine superradiation. *Proceedings of the National Academy of Sciences USA*, 116(16): 7916 – 7925.

Onley D & Scofield P. (2007). Field guide to the Albatrosses, Petrels and Shearwaters of the World. *Christopher Helm*, London.

Otley H, Tansell J & Scofield P. (2017). A comprehensive demographic assessment of the endangered Fiordland crested penguin *Eudyptes pachyrhynchus*. *New Zealand Journal of Zoology*, 44(2): 144 – 162.

Otley H, Edmonds H, Hiscock J, Newton G, Tansell J, van Klink P, Wilson R & Westbrooke I. (2018). Assessing the population trend and threats to New Zealand's Fiordland crested penguin using counting and demographic modelling approaches. *New Zealand Journal of Ecology*, 42(2): 125 – 136.

Overeem RL, Peucker AJ, Austin CM, Dann P & Burridge CP. (2008). Contrasting genetic structuring between colonies of the World's smallest penguin, *Eudyptula minor* (Aves: Spheniscidae). *Conservation Genetics*, 9(4): 893 – 905.

Pan H, Cole TL, Bi X, Fang M, Zhou C, Yang Z, Ksepka DT, Hart T, Bouzat JL, Argilla LS, Bertelsen MF, Boersma PD, Bost C, Cherel Y, Dann P, Fiddaman SR, Howard P, Labuschagne K, Mattern T, Miller G, Parker P, Phillips RA, Quillfeldt P, Ryan PG, Taylor H, Thompson DR, Young MJ, Ellegaard MR, Gilbert MTP, Sindings MS, Pacheco G, Shepherd LD, Tennyson AJD, Grosser S, Kay E, Nupen LJ, Ellenberg U, Houston DM, Hart-Reeve A, Johnson K, Masello JF, Stracke T, McKinlay B, Borboroglu García P, Zhang DX & Zhang G. (in press). High-coverage genomes to elucidate the evolution of penguins. *GigaScience.*

Parent CE & Crespi BJ. (2006). Sequential colonization and diversification of Galápagos endemic land snail genus *Bulimulus* (Gastropoda, Stylommatophora). *Evolution*, 60(11): 2311 – 2328.

Parent CE, Cannone A & Petren K. (2008). Colonization and diversification of Galápagos terrestrial fauna: a phylogenetic and biogeographical synthesis. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 363(1508): 3347 – 3361.

Park T & Fitzgerald EMG. (2012). A review of Australian fossil penguins (Aves: Sphenisciformes). *Memoirs of Museum Victoria*, 69: 309 – 325.

Park T, Fitzgerald EMG, Gallagher SJ, Tomkins E & Allan T. (2016). New Miocene Fossils and the History of Penguins in Australia. *Plos One*, 11(4): e0153915.

Parmesan C & Yohe G. (2003). A globally coherent fingerprint of climate change impacts across natural systems. *Nature*, 421(6918): 37 – 42.

Patten MA & Unitt P. (2002). Diagnosability versus mean differences of Sage Sparrow subspecies. *The Auk*, 119(1): 26 – 35.

Paulay G. (1994). Biodiversity on Oceanic Islands: Its Origin and Extinction. *American Zoologist*, 34(1): 134 – 144.

Pembleton LW, Cogan NOI & Forster JW. (2013). StAMPP: an R package for calculation of genetic differentiation and structure of mixed-ploidy level populations. *Molecular Ecology Resources*, 13(5): 946 – 952.

Pena F, Poulin E, Dantas GP, Gonzalez-Acuna D, Petry MV & Vianna JA. (2014). Have historical climate changes affected Gentoo penguin (*Pygoscelis papua*) populations in Antarctica? *PLoS One*, 9(4): e95375.

Peucker AJ, Dann P & Burridge CP. (2009). Range-wide phylogeography of the little penguin (*Eudyptula minor*): evidence of long-distance dispersal. *The Auk*, 126(2): 397 – 408.

Pimm L, Alibhai S, Berg R, Dehgan A, Giri C, Jewell Z, Joppa L, Kays R & Loarie S. (2015). Emerging technologies to conserve biodiversity. *Trends in Ecology and Evolution*, 30(11): 685 – 696.

Pina-Martins F, Silva DN, Fino J & Paulo OS. (2017). Structure_threader: An improved method for automation and parallelization of programs structure, fastStructure and MavericK on multicore CPU systems. *Molecular Ecology Resources*, 17(6): e268 – e274.

Powlesland RG. (1984). Seabirds found dead on New Zealand beaches in 1982 and a review of penguin recoveries since 1960. *Notornis*, 31: 155 – 171.

Pritchard JK, Stephens M & Donnelly P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, 155(2): 945 – 959.

Prost S & Anderson CNK. (2011). TempNet: a method to display statistical parsimony networks for heterochronous DNA sequence data. *Methods in Ecology and Evolution*, 2(6): 663 – 667.

Prüfer K, Racimo N, Jay F, Sankararaman S, Sawyer A, Renaud G, Sudmant PH, de Filippo C, Li H, Mallick S, Dannemann M, Fu Q, Kircher M, Kuhlwilm M, Lachmann M, Meyer M, Ongyerth M, Siebauer M, Theunert C, Tandon A, Moorjani P, Pickrell J, Mullikin JC, Vohr SH, Green RE, Hellmann I, Johnson PLF, Blanche H, Cann H, Kitzman JO, Shendure J, Eichler EE, Lein ES, Bakken TE, Golovanova LV, Doronichev VB, Shunkov MV, Derevianko AP, Viola B, Slatkin M, Reich D, Kelso J & Pääbo S. (2014). The complete genome sequence of a Neanderthal from the Altai Mountains. *Nature*, 545(7481): 43.

Pütz K, Rey AR & Otley H. (2013). In Penguins: Natural History and Conservation. *University of Washington Press*, Washington.

Raine JI, Beu AG, Boyes AF, Campbell HJ, Cooper RA, Crampton JS, Crundwell MP, Hollis CJ & Morgans HEG. (2015). Revised calibration of the New Zealand Geological Timescale: NZGT2015/1. GNS Science report 2012/39 (GNS Science, Lower Hutt NZ) 53 p.

Rains D, Weimerskirch H & Burg TM. (2011). Piecing together the global population puzzle of Wandering Albatrosses: Genetic analysis of the Amsterdam Albatross *Diomedea amsterdamensis*. *Journal of Avian Biology*, 42(1): 69 – 79.

Rainsley E, Turney CS, Golledge NR, Wilmshurst JM, McGlone MS, Hogg AG, Li B, Thomas ZA, Roberts R, Jones RT & Palmer JG. (2019). Pleistocene glacial history of the New Zealand subantarctic islands. *Climate of the Past*, 15(2): 423-448.

Raman H, Raman R, Kilian A, Detering F, Carling J, Coombes N, Diffey S, Kadkol D, McCully M, Ruperao P, Parkin IAP, Batley J, Luckett DJ & Wratten N. (2014). Genome-wide delineation of natural variation for pod shatter resistence in *Brassica napis*. *PloS One*, 9(7): e101673.

Rambaut A, Drummond AJ & Suchard M. (2013). Tracer v1.6 – MCMC Trace Analysis Package. Institute of Evolutionary Biology, University of Edinburgh, UK.

Ramos B, González-Acuña D, Loyola DE, Johnson WE, Parker PG, Massaro M, Dantes GPM, Miranda MD & Vianna JA. (2018) Landscape genomics: natural selection drives the evolution of mitogenome in penguins. *BMC Genomics*, 19: 53.

Randall RM & Randall BM. (1986). The diet of jackass penguins *Spheniscus demersus* in Algoa Bay, South Africa, and its bearing on population declines elsewhere. *Biological conservation*, 37(2): 119 – 134.

Ratcliffe N & Trathan P. (2012). A review of the diet and at-sea distribution of penguins breeding within the CAMLR Convention Area. *CCAMLR Science*, 19: 75 – 114.

Rawlence NJ, Perry GLW, Smith IWG, Scofield RP, Tennyson AD, Matisso-Smith EA, Boessenkool S, Austin JJ & Waters JM. (2015a). Radiocarbon-dating and ancient DNA reveal rapid replacement of extinct prehistoric penguins. *Quaternary Science Reviews*, 112: 59 – 65.

Rawlence NJ, Kennedy M, Anderson CNK, Prost S, Till CE, Smith IWG, Scofield RP, Tennyson AJD, Hamel J, Lalas C, Matisoo-Smith EA & Waters JM. (2015b). Geographically contrasting biodiversity reductions in a widespread New Zealand seabird. *Molecular Ecology*, 24(18): 4605 – 4616.

Rawlence NJ, Collins CJ, Anderson CNK, Maxwell JJ, Smith IWG, Robertson BC, Knapp M, Horsburgh KA, Stanton J-A L, Scofield RP, Tennyson AJD, Matisoo-Smith EA & Waters JM. (2016). Human-mediated extirpation of the unique Chatham Islands sea lion and implications for the conservation management of remaining New Zealand sea lion populations. *Molecular Ecology*, 25(16): 3950 – 3961.

Rawlence NJ, Kardamaki A, Easton LJ, Tennyson AJD, Scofield RP & Waters JM. (2017a). Ancient DNA and morphometric analysis reveal extinction and replacement of New Zealand's unique black swans. *Proceedings of the Royal Society B: Biological Sciences*, 284(1859): 20170876.

Rawlence NJ, Till CE, Easton LJ, Spencer HG, Schuckard R, Melville DS, Scofield RP, Tennyson AJD, Rayner MJ, Waters JM & Kennedy M. (2017b). Speciation, range contraction and extinction in

the endemic New Zealand king shag complex. *Molecular Phylogenetics and Evolution*, 115: 197 – 209.

Rawlence NJ, Tennyson AJD, Cole TL, Verry A & Scofield RP. (2018). Evidence for breeding of *Megadyptes* penguins in North Island at the time of human arrival. *New Zealand Journal of Zoology*, 46(2): 165 – 173.

Reid BN, Naro-Maciel E, Torres Hahn A, FitzSimmons NN & Gehara M. (2019). Geography best explains global patterns of genetic diversity and postglacial co-expansion in marine turtles. *Molecular Ecology*, 28(14): 3358 – 3370.

Rexer-Huber K. (2017). White-chinned petrel distribution, abundance and connectivity have circumpolar conservation implications. Thesis, Doctor of Philosophy. University of Otago.

Rexer-Huber, Veale AJ, Catry P, Cherel Y, Dutoit L, Foster Y, McEwan JC, Parker GC, Phillips RA, Ryan PG, Stanworth AJ, van Stijn T, Thompson DR, Waters J & Robertson BC. (in press). Genomics detects population structure within and between ocean basins in a circumpolar seabird: the white-chinned petrel. *Molecular Ecology*.

Rheindt FE & Edwards SV. (2011). Genetic introgression: an integral but neglected component of speciation in birds. *The Auk*, 128(4): 620 – 632.

Richards MD. (2019). Two giant penguins from the Eocene-Oligocene of Otago, New Zealand. Thesis, Master of Science. University of Otago.

Ritchie PA, Millar CD, Gibb GC, Baroni C & Lambert DM. (2004). Ancient DNA enables timing of the Pleistocene origin and Holocene expansion of two Adélie Penguin lineages in Antarctica. *Molecular Biology and Evolution*, 21(2): 240 – 248.

Roberts CM, Duncan RP & Wilson K-J. (2007). Burrowing seabirds affect forest regeneration, Rangitira Island, Chatham Islands, New Zealand. *New Zealand Journal of Ecology*, 31(2): 208 – 222.

Robertson HA, Baird K, Dowding JE, Elliott GP, Hitchmough RA, Miskelly CM, McArthur N, O'Donnell CFJ, Sagar PM, Scofield RP & Taylor GA. (2017). Conservation status of New Zealand birds, 2016. New Zealand Threat Classification Series 19. New Zealand Department of Conservation, May 2017.

Robson B, Glass T, Glass N, Glass J, Green J, Repetto C, Rodgers G, Ronconi RA, Ryan PG, Swain G & Cuthert RJ. (2011). Revised population estimate and trends for the Endangered Northern Rockhopper Penguin *Eudyptes moseleyi* at Tristan da Cunha. *Bird Conservation International*, 21(4): 454 – 459.

Rohland N, Siedel H & Hofreiter M. (2010). A rapid column-based ancient DNA extraction method for increased sample throughout. *Molecular Ecology Resources*, 10(4): 677 – 683.

Rohland N, Harney E, Mallick S, Nordenfelt S & Reich D. (2015). Partial uracil-DNA-glycosylase treatment for screening of ancient DNA. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 370(1660): 20130624.

Ropert-Coudert Y, Chiaradia A, Ainley D, Barbosa A, Boersma P D, Brasso R, Dewar M, Ellenberg U, García-Borboroglu P, Emmerson L, Hickcox R, Jenouvrier S, Kato A, Ruth McIntosh R, Lewis P,

Ramírez F, Ruoppolo V, Ryan PG, Seddon PJ, Brain Sherley R, Vanstreels RET, Waller LJ, Woehler EJ & Trathan PN. (2019). Happy Feet in a hostile world? The future of penguins depends on proactive management of current and expected threats. *Frontiers in Marine Science*, 6: 248.

Rutschmann S, Detering H, Simon S, Funk DH, Gattolliat J-L, Hughes SJ, Raposeiro PM, DeSalle R, Sartori M & Monaghan MT. (2017). Colonization and diversification of aquatic insects on three Macronesian archipelagos using 59 nuclear loci derived from a draft genome. *Molecular Phylogenetics and Evolution*, 107: 27 – 38.

Salis AT, Easton LJ, Robertson BC, Gemmell N, Smith IWG, Weisler M, Waters JM & Rawlence NJ. (2016). Myth or relict: does ancient DNA detect the enigmatic upland seal? *Molecular Phylogenetics and Evolution*, 97: 101 – 106.

Salton M, Kliska K, Carmichael N & Alderman R. (2019). Population status of the endemic royal penguin (*Eudyptes schlegeli*) at Macquarie Island. *Polar Biology*, 42(4): 771 – 781.

Sander M, Balbão TC, Costa ES, Dos Santos CR & Petry MV. (2007). Decline of the breeding population of *Pygoscelis antarctica* and *Pygoscelis adeliae* on Penguin Island, South Shetland, Antarctica. *Polar Biology*, 30(5): 651 – 654.

Sansaloni C, Petroli C, Jaccoud D, Carling J, Detering F, Frattapaglia D & Kilian A. (2011). Diversity Arrays Technology (DArT) and next generation sequencing combined: genome-wide, high throughput, highly informative genotyping for molecular breeding of *Eucalyptus*. *BMC Proceedings*, 5(Suppl 7): 54.

Scarlett RJ. (1979). Avifauna and man. In Birds of a feather: Osteological and archaeological papers from the South Pacific in honour of R. J. Scarlett. *New Zealand Archaeological Association Monograph II*. Oxford.

Scasso RA, McArthur JM, Del Rio CJ, Martınez S & Thirlwall MF. (2001). 87Sr/86Sr Late Miocene age of fossil molluscs in the 'Entrerriense' of the Valdés Peninsula (Chubut, Argentina). *Journal of South American Earth* Sciences, 14(3): 319 – 329.

Schlosser JA, Dubach JM, Garner TWJ, Araya B, Bernal M, Simeone A, Smith KA & Wallace RS. (2009). Evidence for gene flow differs from observed dispersal patterns in the Humboldt penguin, *Spheniscus humboldti*. *Conservation Genetics*, 10(4): 839 – 849.

Schubert M, Lindgreen S & Orlando L. (2016). AdapterRemoval v2: rapid adapter trimming, identification, and read merging. *BMC Research Notes*, 12(1): 88.

Scofield P, Worthy TH & Schlumpf H. (2003). What birds were New Zealand's first people eating? – Wairau Bar's avian remains re-examined. *Records of Canterbury Museum*, 17: 17 – 35.

Scofield RP, Mitchell KJ, Wood JR, de Pietri VL, Jarvie S, Llamas B & Cooper A. (2017). The origin and phylogenetic relationships of the New Zealand ravens. *Molecular Phylogenetics and Evolution*, 106(1): 136 – 143.

Scott JM & Turnbull IM. (2019). Geology of New Zealand's Sub-Antarctic Islands. *New Zealand Journal of Geology and Geophysics*. DOI: 10.1080/00288306.2019.1600557.

Seersholm FV, Cole TL, Grealy A, Rawlence NJ, Greig K, Knapp M, Stat M, Hansen AJ, Easton LJ, Shepherd L, Tennyson A, Scofield P, Walter R & Bunce M. (2018). Subsistence practices, past biodiversity, and anthropogenic impacts revealed by New Zealand-wide ancient DNA survey. *Proceedings of the National Academy of Sciences*. 115(30): 7771 – 7776.

Sequeira AS, Sijapati M, Lanteri AA & Roque Albelo L. (2008). Nuclear and mitochondrial sequences confirm complex colonization patterns and clear species boundaries for flightless weevils in the Galápagos archipelago. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 363(1508): 3439 – 3451.

Shaughnessy PD. (1975). Variation in facial colour of the Royal Penguin. *Emu-Austral Ornithology*, 75(3): 147 – 152.

Shaw KL & Gillespie RG. (2016). Comparative phylogeography of oceanic archipelagos: Hotspots for inferences of evolutionary process. *Proceedings of the National Academy of Sciences USA*, 113(29): 7986 – 7993.

Shepherd LD & Lambert DM. (2008). Ancient DNA and conservation: lessons from the endangered kiwi of New Zealand. *Molecular Ecology*, 17(9): 2174 – 2184.

Shepherd LD, Worthy TH, Tennyson AJD, Scofield RP, Ramstad KM & Lambert DM. (2012). Ancient DNA analyses reveal contrasting phylogenetic patterns amongst kiwi (*Apteryx* spp.) and a recently extinct lineage of spotted kiwi. *PLoS one*, 7(8): e42384.

Sherley RB, Barham PJ, Barham BJ, Crawford RJM, Dyer BM, Leshoro TM, Makhado AB, Upfold L & Underhill LG. (2013). Growth and decline of a penguin colony and the influence on nesting density and reproductive success. *Population Ecology*, 56(1): 119 – 128.

Simeone A, Hiriart-Bertrand L, Reyes-Arriagada R, Halpern M, Dubach J, Wallace R, Pütz K, Lüthi B. (2009). Heterospecific pairing and hybridisation between wild Humboldt and magellanic penguins in Southern Chile. *The Condor: Ornithological Applications*, 111(3): 544 – 550.

Simpson GG. (1971a). Fossil penguin from the late Cenozoic of South Africa. *Science*, 171(3976): 1144 – 1145.

Simpson GG. (1971b). Review of fossil penguins from Seymour Island. *Proceedings of the Royal Society of London B: Biological Sciences*. 178(1053): 357 – 387.

Simpson KNG. (2008). The 'Mystery Penguin': an additional Snares Penguin *Eudyptes pachyrhynchus robustus* for Tasmania. *The Tasmanian Naturalist*, 130: 42 – 51.

Sinton WC, Hauff F, Hoernle K & Werner R. (2018). Age progressive volcanism opposite Nazca plate motion: Insights from seamounts on the northeastern margin of the Galapagos Platform. *Lithos*, 310: 342 – 354.

Skoglund P, Northoff BH, Shunkov MV, Derevianko AP, Pääbo S, Krause J & Jakobsson M. 2014. Separating endogenous ancient DNA from modern day contamination in a Siberian Neanderthal. *Proceedings of the National Academy of Sciences USA*, 111(6): 2229 – 2234.

Slack KE, Jones CM, Ando T, Harrison GL, Fordyce RE, Arnason U & Penny D. (2006). Early penguin fossils, plus mitochondrial genomes, calibrate avian evolution. *Molecular Biology and Evolution*, 23(6): 1144 – 1155.

Slon V, Hopfe C, Weiß CL, Mafessoni F, de la Rasilla M, Lalueza-Fox C, Rosas A, Soressi M, Knul MV, Miller R, Stewart JR, Derevianko AP, Jacobs Z, Li B, Roberts RG, Shunkov MV, de Lumley H, Perrenoud C, Gušić I, Kućan Ž, Rudan P, Aximu-Petri A, Essel E, Nagel S, Nickel B, Schmidt A, Prüfer K, Kelso J, Burbano HA, Pääbo S & Meyer M. (2017). Neandertal and Denisovan DNA from Pleistocene sediments. *Science*. 356(6338): 605 – 608.

Smith CI, Tank S, Godsoe W, Levenick J, Strand E, Esque T & Pellmyr O. (2011). Comparative phylogeography of a coevolved community: concerted population expansions in Joshua trees and four yucca moths. *PloS One*, 6(10): e25628.

Spatz DR, Zilliacus KM, Holmes ND Butchart SMH, Genovesi P, Ceballos G, Tershy BR & Croll DA. (2017). Globally threatened vertebrates on islands with invasive species. *Consevation Biology*, 3(10): e1603080.

Stamatakis A. (2006). RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics,* 22(21): 2688 – 2690.

Stamatakis A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics,* 30(9): 1312 – 1313.

Steadman DW. (1989). Extinction of birds in eastern Polynesia: a review of the record, and comparisons with other Pacific island groups. *Journal of Archaeological Science*, 16(2): 177 – 205.

Stewart J, Lister AM, Barnes I & Dalén L. (2010). Refugia revisited: individualistic responses of species in space and time. *Proceedings of the Royal Society of London B: Biological Sciences*, 277(1682): 661 – 671.

Stonehouse B. (1967). The general biology and thermal balances of penguins. In Advances in Ecological Research. *Academic Press*.

Stuart YE, Losos JB & Algar AC. (2012). The island-mainland species turnover relationship. *Proceedings of the Royal Society B: Biological Sciences*, 279(1744): 4071 – 4077.

Subramanian S, Beans-Picón G, Swaminathan SK, Millar CD & Lambert DM. (2013). Evidence for a recent origin of penguins. *Biology Letters*, 9(6): 20130748.

Taylor GA. (2000). Action Plan for Seabird Conservation in New Zealand. Part A: Threatened Seabirds. Threatened Species Occasional Publication No. 16. Wellington, Department of Conservation.

Tennyson AJD & Millener PR. (1994). Bird extinctions and fossil bones from Mangere Island, Chatham Islands. *Notornis* supplement, 41: 165 – 178.

Tennyson A & Martinson P. (2007). Extinct Birds of New Zealand Revised Edition. *Te Papa Press*, Wellington.

Thatje S, Hillenbrand CD, Mackensen A & Larter R. (2008). Life hung by a thread: endurance of Antarctic fauna in glacial periods. *Ecology*, 89(3): 682 – 692.

Thiébot JB, Cherel Y, Trathan PN & Bost CA. (2011). Inter-populatoin segregation in the wintering areas of macaroni penguins. *Marine Ecology Progress Series*, 421: 279 – 290.

Thiébot JB, Cherel Y, Trathan PN & Bost CA. (2012). Coexistence of oceanic predators on wintering areas explained by population-scale foraging segregation in space or time. *Ecology*, 93(1): 12 – 130.

Thomas DB & Ksepka DT. (2013). A history of shifting fortunes for African penguins. *Zoological Journal of the Linnean Society*, 168(1): 207 – 219.

Thompson DWJ & Solomon S. (2002). Interpretation of recent Southern Hemisphere climate change. *Science*, 296(5569): 895 – 899.

Thomson VA, Lebrasseur O, Austin JJ, Hunt TL, Burney DA, Denham T, Rawlence NJ, Wood JR, Gongora J, Girland Flink L, Linderholm A, Dobney K, Larson G & Cooper A. (2014). Using ancient DNA to study the origins and dispersal of ancestral Polynesian chickens across the Pacific. *Proceedings of the National Academy of Sciences USA*, 111(13): 4826 – 4831.

Tizard J, Patel S, Waugh J, Tavares E, Bergmann T, Gill B, Norman J, Christidis L, Scofield P, Haddrath O, Baker A, Lamert D & Millar C. (2019). DNA barcoding a unique avifauna: an important tool for evolution, systematics and conservation. *BMC Evolutionary Biology*, 19(1): 52.

Towns DR & Ballantine WJ. (1993). Conservation and restoration of New Zealand island ecosystems. *Trends in Ecology and Evolution*, 8(12): 452 – 457.

Trakhenbrot A, Nathan R, Perry G & Richardson DM. (2005). The importance of long-distance dispersal in biodiversity conservation. *Diversity and Distributions*, 11(2): 173 – 181.

Travers HH & Travers WTL. (1872). On the birds of the Chatham Islands, with introductory remarks on the avi-fauna and flora of the islands in their relation to those of New Zealand. *Transactions and proceedings of the New Zealand Institute*, 5(1): 212 – 222.

Trewick SA. (2000). Molecular evidence for dispersal rather than vicariance as the origin of flightless insect species on the Chatham Islands, New Zealand. *Journal of Biogeography*, 27(5): 1189 – 1200.

Trewick SA & Worthy TH. (2001). Origins and prehistoric ecology of Takahe based on morphometric, molecular, and fossil data. In The Takahe: Fifty Years of Conservation Management and Research. *University of Otago Press*, Dunedin.

Trivelpiece WZ, Hinke JT, Miller AK, Reiss CS, Trivelpiece SG & Watters GM. (2011). Variability in krill biomass links harvesting and climate warming to penguin population changes in Antarctica. *Proceedings of the National Academy of Sciences USA*, 108(18): 7625 – 7628.

Trucchi E, Gratton P, Whittington JD, Cristofari R, Le Maho Y, Stenseth NC & Le Bohec C. (2016). King penguin demography since the last glaciation inferred from genome-wide data. *Proceedings of the Royal Society B: Biological Sciences*, 281(1787): 20140528.

Udvardy MDF. (1987). The biogeographical realm Antarctica: a proposal. *Journal of the Royal Society of New Zealand*. 17(2): 187 – 194.

Valente L, Etienne RS & Rarcia-R JC. (2019). Deep macroevolutionary impact of humans on New Zealand's unique avifauna. *Current Biology*, 19(15): 2563 – 2569.

van Tets GF & O'Connor S. (1983). The Hunter Island Penguin, an extinct new genus and species from a Tasmanian midden. *Records of the Queen Victoria Museum*, 81: 1 – 11.

Vianna JA., Noll D, Dantas GPM, Petry MV, Barbosa A, González-Acuña D, Le Bohec C, Bonadonna F & Poulin E. (2017) Marked phylogeographic structure of Gentoo penguin reveals an ongoing diversification process along the Southern Ocean. *Molecular Phylogenetics and Evolution*, 107: 486 – 498.

Walsh SA & Suárez ME. (2006). New penguin remains from the Pliocene of northern Chile. *Historical Biology*, 18(2):119 – 130.

Walther GR, Post E, Menzel A, Parmesan C, Beebee TJC, Fromentin J-M, Hoegh-Guldberg O & Bairlein F. (2002). Ecological responses to recent climate change. *Nature*, 416(6879): 389.

Waltoft BL & Hobolth A. (2018). Non-parametric estimation of population size changes from the site frequency spectrum. *Statistical Applications in Genetics and Molecular Biology*, 17(3).

Wanner H, Beer J, Bütikofer J, Crowley TJ, Cubasch U, Flückiger J, Goosse H, Grosjean M, Fortunat J, Kaplan JO, Küttel M, Müller SA, Prentice C, Solomin O, Stocker TF, Tarasov P, Wagner M & Widmann M. (2008). Mid- to Late Holocene climate change: an overview. *Quaternary Science Reviews*, 27(19 – 20): 1791 – 1828.

Warham J. (1974a). The breeding biology and behaviour of the snares crested penguin. *Journal of the Royal Society of New Zealand*, 4:63 – 108

Warham J. (1974b). The Fiordland crested penguin *Eudyptes pachyrhynchus*. *The Ibis*, 116: 1 – 27.

Warham J. (1975). The crested penguins. The biology of penguins, pp 189 – 269.

Warnock RCM, Parham JF, Joyce WG, Lyson TR & Donoghue PCJ. (2015). Calibration uncertainty in molecular dating analyses: there is no substitute for the prior evaluation of time priors. *Proceedings of the Royal Society B: Biological Sciences*, 282(1798): 20141013.

Waters JM, Fraser CI & Hewitt GM. (2013). Founder takes all: density-dependent processes structure biodiversity. *Trends in Ecology and Evolution*, 28(2): 78 – 85.

Waters JM & Grosser S. (2016). Managing shifting species: ancient DNA reveals conservation conundrums in a dynamic world. *Bioessays*, 38(11): 1177 – 1184.

Waters JM, Fraser CI, Maxwell JJ & Rawlence NJ. (2017). Did interaction between human pressure and Little Ice Age drive biological turnover in New Zealand? *Journal of Biogeography*, 44(7): 1481 – 1490.

Wheeler TJ & Eddy SR. (2013). nhmmer: DNA homology search with profile HMMs. *Bioinformatics*, 29(19): 2487 – 2489.

White RW & Clausen AP. (2002). Rockhopper *Eudyptes chrysocome chrysocome* x macaroni *E. chrysolophus* penguin hybrids apparently breeding in the Falkland Islands. *Marine Ornithology*, 30: 40 – 42.

Whittaker RJ & Fernández-Palacios JM. (2007). Island biogeography: ecology, evolution, and conservation. *Oxford University Press*, Oxford.

Wilmshurst JM, Anderson AJ, Higham TFG & Worthy TH. (2008). Dating the late prehistoric dispersal of Polynesians to New Zealand using the commensal Pacific rat. *Proceedings of the National Academy of Sciences USA*, 105(22): 7676 – 7680.

Wilmshurst JM, Hunt TL, Lipo CP & Anderson AJ. (2011). High-precision radiocarbon dating shows recent and rapid initial human colonization of East Polynesia. *Proceedings of the National Academy of Sciences USA*, 108(5): 1815 – 1820.

Wilson JT. (1963). A possible origin of the Hawaiian Islands. *Canadian Journal of Physics*, 41(6): 863 – 870.

Woehler E & Gilbert C. (1990). Hybrid rockhopper-macaroni penguins, interbreeding and mixed species pairs at Heard and Marion Islands. *Emu-Austral Ornithology*, 90(3): 198 – 201.

Woehler E. (1992). Records of vagrant penguins from Tasmania. *Marine Ornithology*, 20: 61 – 73.

Woehler EJ. (1995). Bill morphology of Royal and Macaroni Penguins, and geographic variation within eudyptid penguins. The penguins. *Surrey Beatty & Sons Pty Ltd*, Chipping Norton.

Woehler EJ, Cooper J, Croxall JP, Fraser WR, Kooyman GL, Miller GD, Nel DC, Patterson DL, Peter H-U, Ribic CA, Salwicka K, Trivelpiece WZ & Weimerskirch H. (2011). A statistical assessment of the status and trends of Antarctic and sub-Antarctic seabirds. *Scientific Committee on Antarctic Research*, Cambridge.

Wood JR, Mitchell KJ, Scofield RP, Tennyson AJD, Fidler AE, Wilmshurst JM, Llamas B & Cooper A. (2014). An extinct nesteorid parrot (Aves, Psittaciformes, Nestoridae) from the Chatham Islands, New Zealand. *Zoological Journal of the Linnean Society*, 172(1): 185 – 199.

Wood JR, Alcover JA, Blackburn TM, Bover P, Duncan RP, Hume JP, Louys J, Meijer HJM, Rando JC & Wilmshurst JM. (2017). Island extinctions: processes, patterns and potential for ecosystem restoration. *Environmental Conservation*, 44(4): 348 – 358.

Worthy TH. (1997). The identification of fossil *Eudyptes* and *Megadyptes* bones at Marfells Beach, Marlborough, South Island. *New Zealand Natural Sciences*, 23: 71 – 85.

Worthy TH. (1998). A remarkable fossil and archaeological fauna from Marfell's Beach, Lake Grassmere, South Island, New Zealand. *Records of the Canterbury Museum*, 12: 79 – 176.

Worthy TH. (1999). What was on the menu- avian extinction in New Zealand. *New Zealand Journal of Archaeology*, 19: 125 – 160.

Worthy TH & Holdaway RN. (2002). The Lost World of the Moa, Prehistoric Life of New Zealand. *Canterbury University Press*, Christchurch.

Wu J. (2006). Landscape ecology, cross-disciplinarity, and sustainable science. *Landscape Ecology*, 21(1): 1 – 4.

Xie Y, Wu G, Tang J, Luo R, Patterson J, Liu S, Huang W, He G, Gu S, Li S, Zhou X, Lam TW, Li Y, Xu X, Wong GK & Wang J. (2014). SOAPdenovo-Trans: de novo transciptome assembly with short RNA-seq reads. *Bioinformatics*, 30(12): 1660 – 1666.

Xue AT & Hickerson MJ. (2017). Multi-dice: r package for comparative population genomic inference under hierarchical co-demographic models of independent single-population size changes. *Molelcular Ecology Resources*, 17(6): e212 – e224.

Younger JL, Clucas GV, Kooyman G, Wienecke B, Rogers AD, Trathan P, Hart T & Miller KJ. (2015a). Too much of a good thing; sea ice extent may have forced emperor penguins into refugia during the last glacial maximum. *Global Change Biology*, 21(6): 2215 – 2226.

Younger J, Emmerson L, Southwell C, Lelliott P & Miller K. (2015b). Prolifeation of East Antarctic Adélie penguins in response to historical deglaciation. *BMC Evolutionary Biology*, 15(1): 236.

Younger JL, Emmerson LM & Miller KJ. (2015c). The influence of historical climate changes on Southern Ocean marine predator populations: a comparative analysis. *Global Change Biology*, 22(2): 474 – 493.

Younger JL, van den Hoff J, Wienecke B, Hindell M & Miller KJ. (2016). Contrasting responses to climate regime change by sympatric, ice-dependent predators. *BMC Evolutionary Biology*, 16(1): 61.

Younger JL, Clucas GV, Kao D, Rogers AD, Gharbi K, Hart T & Miller KJ. (2017). The challenges of detecting subtle population structure and its importance for the conservation of Emperor penguins. *Molecular Ecology*, 26(15): 3883 – 3897.

Zhang G, Li C, Li Q, Larkin DM, Lee C, Storz JF, Antunes A, Greenwold MJ, Meredith RW, Ödeen A, Cui J, Zhou Q, Xu L, Pan H, Wang Z, Jin L, Zhang P, Hu H, Yang W, Hu J, Xiao J, Yang Z, Liu Y, Xie Q, Yu H, Lian J, Wen P, Zhang F, Li H, Zeng Y, Xiong Z, Lui S, Zhou L, Huang Z, An N, Wang J, Zheng Q, Xiong Y, Wang G, Wang B, Wang J, Fan Y, da Fonseca RR, Alfaro-Núñez A, Schubert M, Orlando L, Mourier T, Howard JT, Ganapathy G, Pfenning A, Whitney O, Rivas MV, Hara E, Smith J, Farré M, Narayan J, Slavov G, Romanov MN, Borges R, Machado JP, Khan I, Springer MS, Gatesy J, Hoffmann FG, Opazo JC, Håstad O, Sawyer RH, Kim H, Kim KW, Kim HJ, Cho S, Li N, Huang Y, Bruford MW, Zhan X, Dixon A, Bertelsen MF, Derryberry E, Warren W, Wilson RK, Li S, Ray DA, Green RE, O'Brien SJ, Griffin D, Johnson WE, Haussler D, Ryder OA, Willerslev E, Graves GR, Alström P, Fjeldså J, Mindell DP, Edwards SV, Braun EL, Rahbek C, Burt DW, Houde P, Zhang Y, Yang H, Wang J, Avian Genome Consortium, Jarvis ED, Gilbert MT & Wang J. (2014a). Comparative genomics reveals insights into avian genome evolution and adaptation. *Science*, 346(6215): 1311 – 1320.

Zhang G, Jarvis ED & Gilbert MTP. (2014b). A flock of genomes. *Science*, 346(6215): 1308 – 1309.

Zhang G, Li B, Li C, Gilbert MT, Jarvis ED, Wang J & Avian Genome Consortium. (2014c). Comparative genomic data of the Avian Phylogenomics Project. *Gigascience*, 3(1): 26.

Zhang G. (2015). Bird sequencing project takes off. *Nature*, 522(7554): 34.

Zink RM & Barrowclough GF. (2008). Mitochondrial DNA under siege in avian phylogeography. *Molecular Ecology*, 17(9), 2107 – 2121.

Zusi RL. (1975). An interpretation of skull structure in penguins; in The Biology of Penguins. *University Park Press*, Baltimore.

Appendix

Additional articles

The following list includes eleven additional articles that I have either led or been a co-author of during my doctoral studies. While these articles are not included as part of my thesis, they are all related to my thesis by the general aims, questions and methodologies.

Pan H*, Cole TL*, Bi X, Fang M, Zhou C, Yang Z, Ksepka DT, Hart T, Bouzat JL, Argilla LS, Bertelsen MF, Boersma PD, Bost C, Cherel Y, Dann P, Fiddaman SR, Howard P, Labuschagne K, Mattern T, Miller G, Parker P, Phillips RA, Quillfeldt P, Ryan PG, Taylor H, Thompson DR, Young MJ, Ellegaard MR, Gilbert MTP, Sinding MS, Pacheco G, Shepherd LD, Tennyson AJD, Grosser S, Kay E, Nupen JL, Ellenberg U, Houston DM, Hart-Reeve A, Johnson K, Masello JF, Stracke T, McKinlay B, García Borboroglu P, Zhang DX, Zhang G. (in press). High-coverage genomes to elucidate the evolution of penguins. *GigaScience*.*Joint first author.

Frugone M-J, López ME, Segovia N, Cole TL, Lowther A, Pistorius P, Dantas G, Petry MV, Bonadonna F, Trathan P, Polanowski A, Wienecke B, Bi K, Wang-Claypool C, Waters JM, Bowie RCK, Poulin E & Vianna JA. (2019). More than the eye can see: Genomic insights into the drivers of genetic differentiation in Royal/Macaroni penguins across the Southern Ocean. *Molecular Phylogenetics and Evolution.* 139: 106563.

Masello JF, Quillfeldt P, Sandoval-Castellano E, Alderman R, Calderón L, Cherel Y, Cole TL, Cuthbert RJ, Marin M, Massaro M, Navarro J, Phillips RA, Ryan PG, Shepherd LD, Suazo CG, Weimerskirch H, Moodley Y. (2019). Additive traits lead to feeding advantage and reproductive isolation, promoting homoploid hybrid speciation. *Molecular Biology and Evolution*. 36: 1671 – 1685.

Díaz FP, Latorre C, Carrasco-Puga G, Wood JR, Wilmshrust JM, Soto DC, Cole TL & Gutiérrez RA. (2019). Multiscale climate change impacts on plant diversity in the Atacama Desert. *Global Change Biology*, 25(5): 1733 – 1745.

Heenan PB, Wood JR & Cole TL. (2018). A partial cpDNA trnL sequence from the extinct legume *Streblorrhiza speciose* confirms its placement in the tribe Coluteae (Fabaceae). *Phytotaxa*, 374(1): 87 – 91.

Rawlence NJ, Tennyson AJD, Cole TL, Verry A & Scofield RP. (2018). Evidence for breeding of *Megadyptes* penguins in North Island at the time of human arrival. *New Zealand Journal of Zoology*, 46(2): 165 – 173.

Seersholm FV, Cole TL, Grealy A, Rawlence NJ, Greig K, Knapp M, Stat M, Hansen AJ, Easton LJ, Shepherd L, Tennyson A, Scofield P, Walter R & Bunce M. (2018). Subsistence practices, past biodiversity, and anthropogenic impacts revealed by New Zealand-wide ancient DNA survey. *Proceedings of the National Academy of Sciences*, 115(30): 7771 – 7776.

Cole TL & Wood JR. (2017). The ancient DNA revolution: the latest era in unearthing New Zealand's faunal history. *New Zealand Journal of Zoology*, 45(2): 91 – 120.

Cole TL & Wood JR. (2017). Hybridisation in the last remaining individuals of the extinct Fiordland population of brown teal (*Anas chlorotis*). *Records of the Canterbury Museum*, 31(1): 159 – 165.

Cole TL, Hammer MP, Unmack PJ, Teske P, Adams M & Beheregaray LB. (2016). Range-wide fragmentation in a threatened fish associated with post-European settlement modification in the Murray-Darling Basin, Australia. *Conservation Genetics*, 17(6): 1377 – 1391.

Wood JR, Crown A, Cole TL & Wilmshurst JM. (2016) Microscopic and ancient DNA profiling of prehistoric dog (kurī) coprolites from northern New Zealand. *Journal of Archaeological Science Reviews*, 6(1): 496-505.